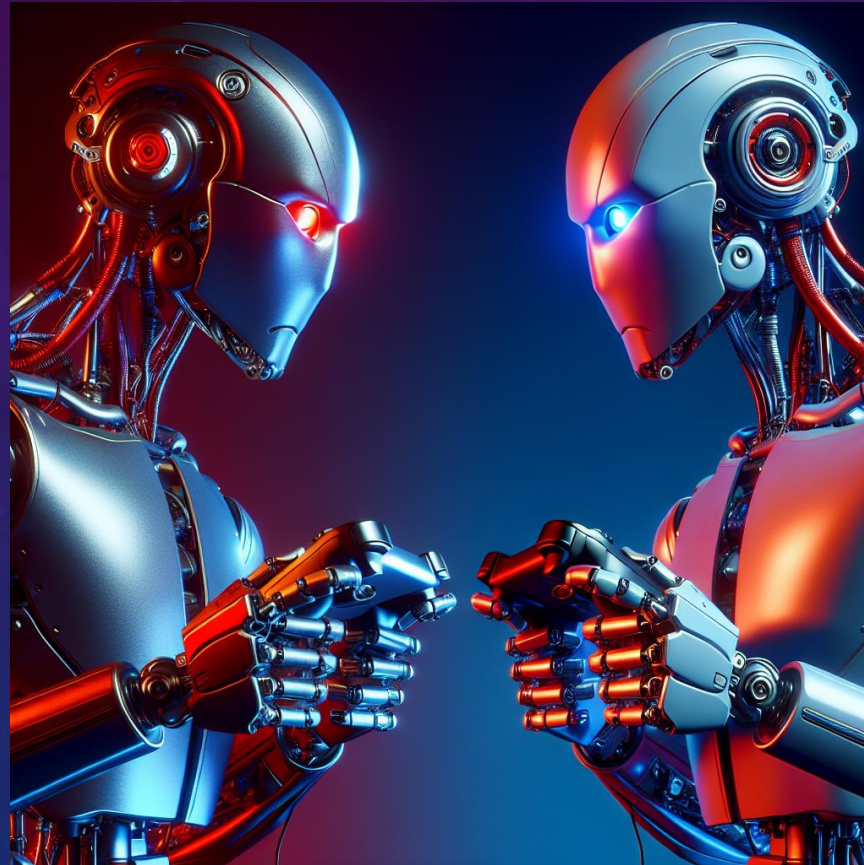


AI UNDERCOVER

AI BOTS IN THE WORLD OF AI CHEAT DETECTION



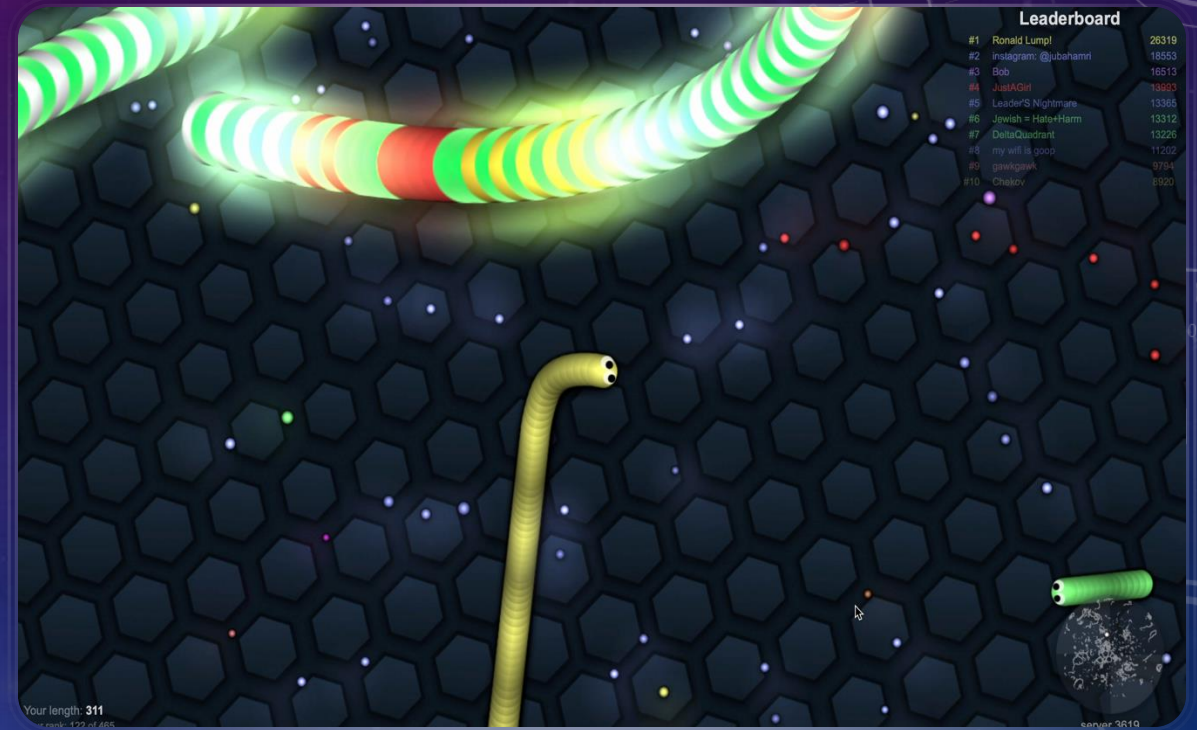
Faraz Akbarzadeh - Santusht Arora - Adil Guluzade
EECS 4461

The background is a gradient of dark blue and purple. It features several faint, white, concentric circles and arcs. A prominent circular scale with degree markings (40, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260) is visible on the left side. Other smaller circles and arcs are scattered across the image, some with arrows indicating direction.

PROJECT OVERVIEW

MEDIA ECOSYSTEM

- Simulates an online ecosystem similar to slither.io
- Players control an Avatar resembling a snake, which consumes multi-colored pellets to grow
- Objective: Grow the longest snake in the server
- AI bots can exploit the game
- Cheat detectors try to identify & eliminate AI cheat bots



AI AGENT ROLES



Humans: Represent regular players making autonomous decisions.



Cheat Detectors: AI agents tasked with identifying and mitigating cheating behaviors.



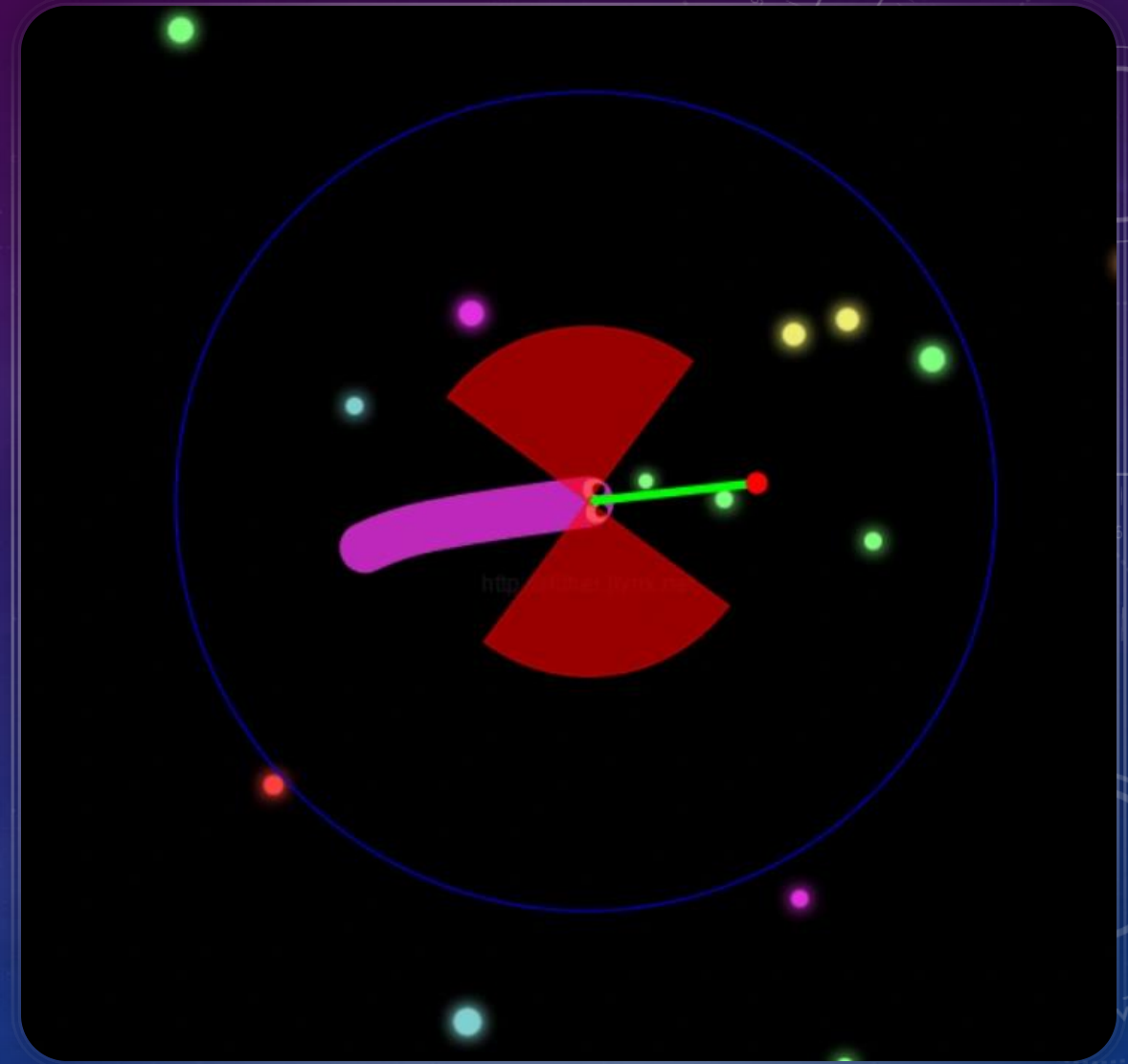
Cheat Bots: Automated agents with enhanced capabilities for cheating.



Cheat Bots: Automated agents flagged red upon detection.

PHENOMENON BEING SIMULATED

- AI bots frequently bombard the game, using different tactics to get a gameplay advantage.
- Because of this, they grow faster and remain safer by preventing game overs.
- Our AI cheat bots track the rate of growth of all agents *, and uses that as a parameter to predict which of the agents might be AI bots



DEMO VIDEO

Boid Flocking Model

Controls

Play Interval (ms)

Render Interval (steps)

☐ Use Threads

RESET

▶

STEP

Model Parameters

Random Seed
42

Number of boids

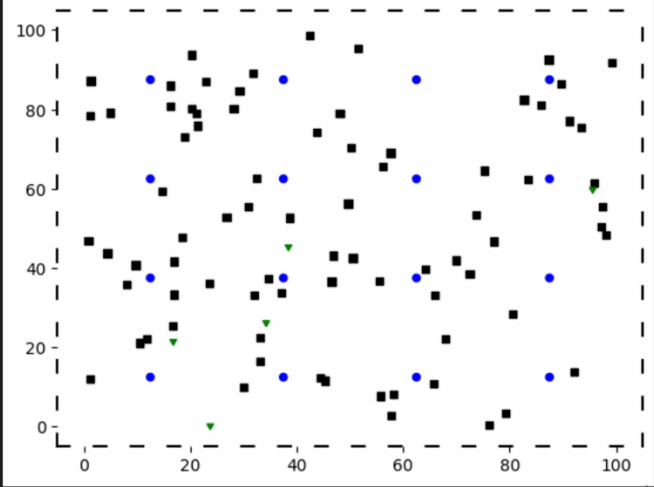
Speed of Boids

Vision of Bird (radius)

Minimum Separation

Information

Step: 0



Simulation Statistics

Eliminated Agents

Eliminated Regular Players: 0

Eliminated Cheaters: 0

Current Agents

Current Regular Players: 79

Current Cheat Detectors: 16

Current Cheaters: 5

Flagged Cheaters: 0

Score Statistics

Regular Player Highest Score: 0

Regular Player Lowest Score: 0

Cheat Bot Highest Score: 0

Cheat Bot Lowest Score: 0

Detection Statistics (Current Step)

True Positives: 0

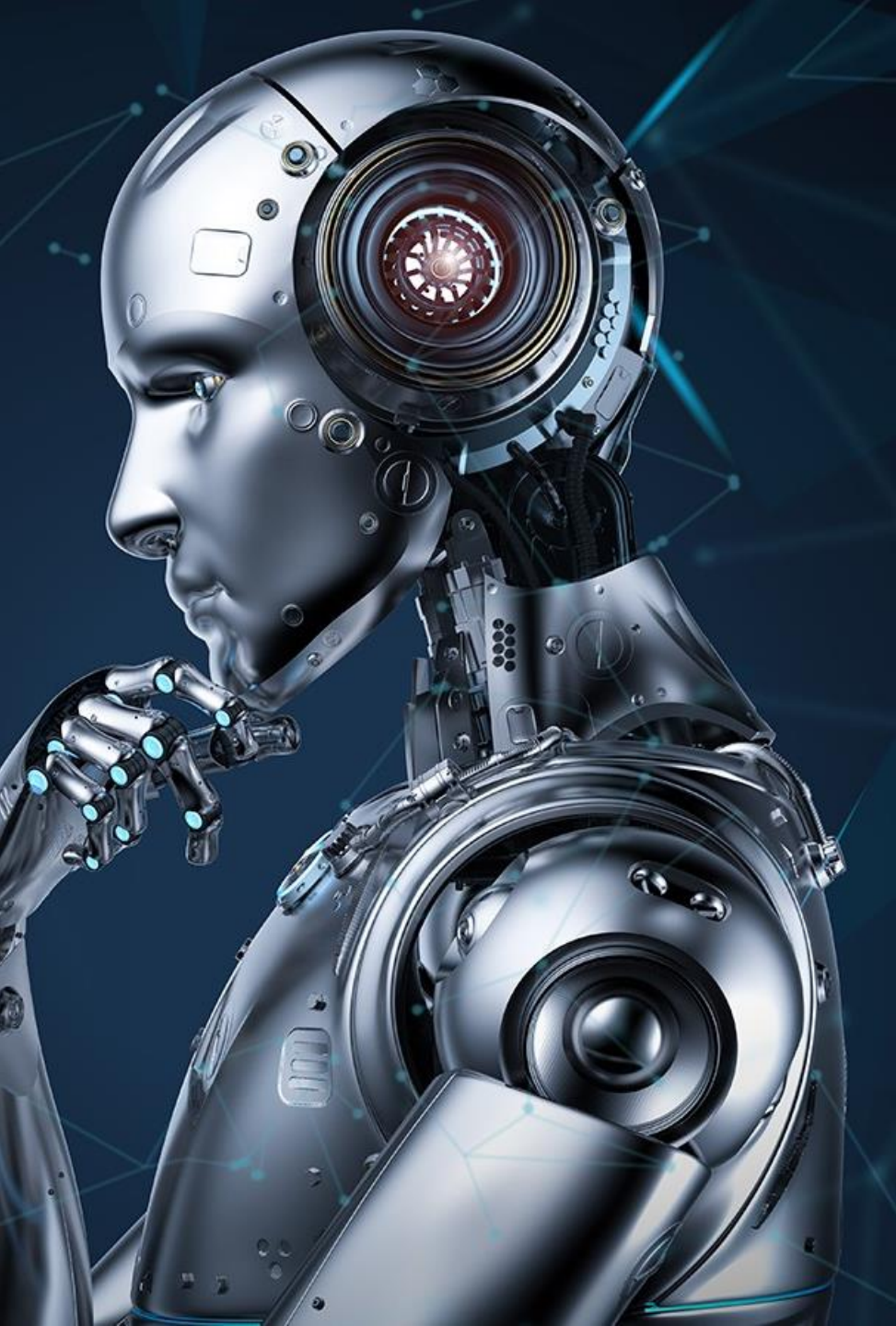
False Positives: 0

True Negatives: 0

False Negatives: 0

The background is a gradient of dark blue and purple, featuring faint, stylized circular patterns and degree markings (40, 150, 160, 170, 180, 190, 200, 220, 230, 240, 250, 260) that suggest a circular or orbital theme.

FINDINGS & INSIGHTS



AI-TO-AI INTERACTION

AI Cheat Bots:

- Distinct AI bots have different ways to grow rapidly based on unfair advantages

Detection Process:

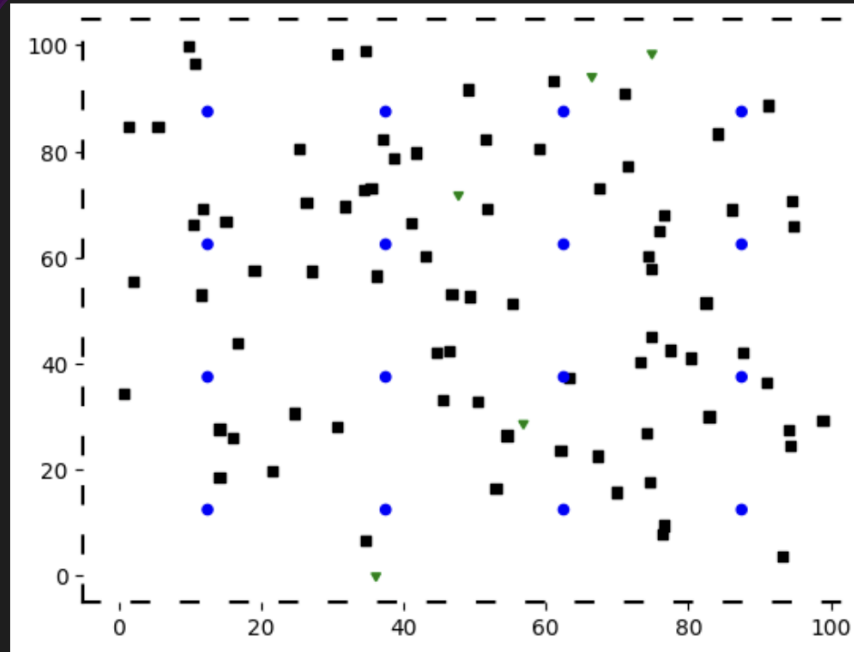
- Blue agents monitor nearby agents
- They assess the rate of growth (score increase) as an indicator of abnormal behavior

Flagging & Removal:

- Agents with unusually high growth are flagged (turned red)
- Flagged agents are eventually removed if they continue to “grow” rapidly

KEY FINDINGS

- Beginning
 - We start with a distribution of randomly placed players
 - Agents are either black, green, or blue
 - All blue agents (cheat detectors) are spread out uniformly



Simulation Statistics

Eliminated Agents

Eliminated Regular Players: 0

Eliminated Cheaters: 0

Current Agents

Current Regular Players: 79

Current Cheat Detectors: 16

Current Cheaters: 5

Flagged Cheaters: 0

Score Statistics

Regular Player Highest Score: 0

Regular Player Lowest Score: 0

Cheat Bot Highest Score: 0

Cheat Bot Lowest Score: 0

Detection Statistics (Current Step)

True Positives: 0

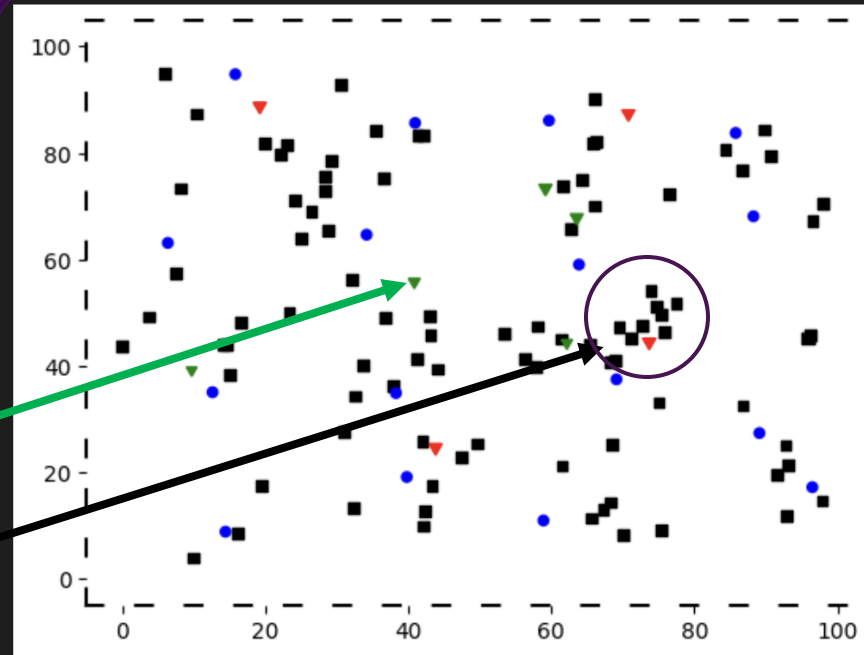
False Positives: 0

True Negatives: 0

False Negatives: 0

KEY FINDINGS

- Running observations
 - Green players start getting flagged (turn red), indicating detection
 - We start to see flocks of human players (black colored agents)
 - Green agents try to blend in with these clusters
 - Green Agents seem to be growing faster than black agents



Simulation Statistics

Eliminated Agents

Eliminated Regular Players: 7
Eliminated Cheaters: 0

Current Agents

Current Regular Players: 92
Current Cheat Detectors: 16
Current Cheaters: 9
Flagged Cheaters: 4

Score Statistics

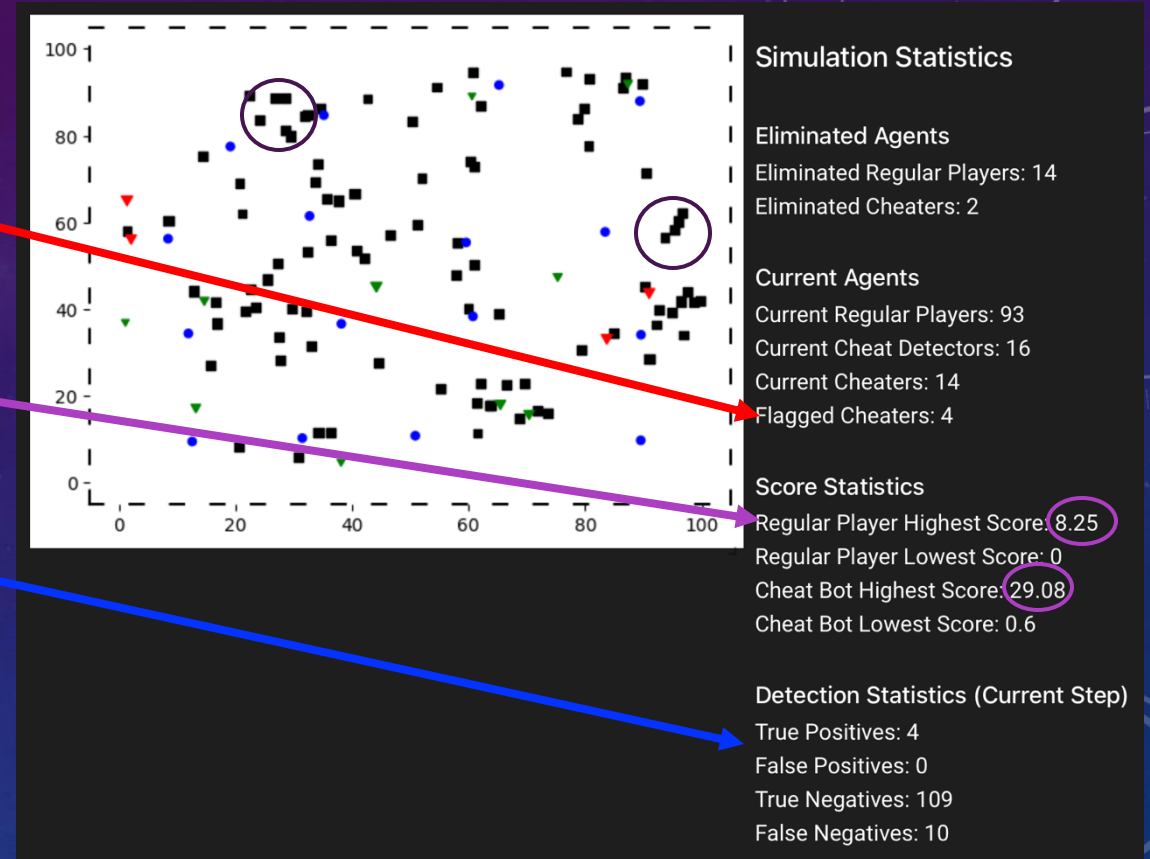
Regular Player Highest Score: 6.03
Regular Player Lowest Score: 0.1
Cheat Bot Highest Score: 22.24
Cheat Bot Lowest Score: 7.0

Detection Statistics (Current Step)

True Positives: 4
False Positives: 0
True Negatives: 108
False Negatives: 5

KEY FINDINGS

- Phase 1 (AI Cheat Bot dominance)
 - The cheat detector is working to identify and eliminate AI
 - Threshold isn't reached yet
 - Human Players still catching up
 - Very few to no false positives
 - A high number of false negatives
 - Some amount of noticeable flocking

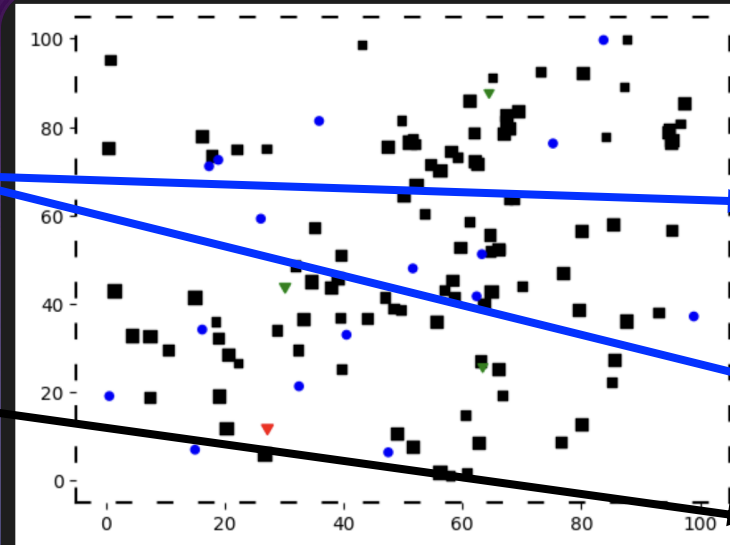


This is the beginning phase where the cheat detection system has already kicked in but the results will show themselves in the upcoming steps.

KEY FINDINGS

- Phase 2 (AI Cheat bot recession)
 - The cheat detector has successfully started eliminating most AI bots
 - Threshold has been breached and triggered
 - Human Players have higher scores than AI agents
 - Rare occurrences of some false positives, but not eliminated, only flagged.
 - Lower number of false negatives
 - High flocking observed

The Cheat bots have not bought an equilibrium to the map, where they are successfully able to handle the cheat bots so that they do not overtake the map and impact human player gameplay. Some AI agents are still there because of the new agents spawning, but the AI makes sure they can not interrupt the game results – very similar to how an actual game ecosystem would work



Simulation Statistics

Eliminated Agents

Eliminated Regular Players: 88

Eliminated Cheaters: 26

Current Agents

Current Regular Players: 105

Current Cheat Detectors: 16

Current Cheaters: 3

Flagged Cheaters: 0

Score Statistics

Regular Player Highest Score: 28.76

Regular Player Lowest Score: 0

Cheat Bot Highest Score: 25.24

Cheat Bot Lowest Score: 13.49

Detection Statistics (Current Step)

True Positives: 0

False Positives: 1

True Negatives: 120

False Negatives: 3

REAL WORLD APPLICATION

- **Online Gaming:**
 - Enhanced cheat detection to preserve game fairness
- **Cybersecurity:**
 - Adaptive defense strategies against evolving AI threats
- **Digital Trust:**
 - Balancing security measures to ensure user trust
- **Content Moderation:**
 - Applications in filtering and managing digital content across platforms



Reflections & ethical considerations

- Figuring out how to remove players
- Figuring out how to flag agents, so that we can identify AI without being unfair to human agents was both a technical and an ethical dilemma
- Keeping and reflecting scores without making actual trails for snakes
- Picking right statistics

Q
Q

&
&

A
A