SEED // SEARCH FOR EXTRAORDINARY EXPERIENCES DIVISION
seed.ea.com

Augmenting Automated Game Testing with Deep Reinforcement Learning

Joakim Bergdahl, <u>Camilo Gordillo</u>, Konrad Tollmar and Linus Gisslen

SEED – Electronic Arts

MOTIVATION

◈ Games are growing in both size and lifetime

◈ We have reached the limit to what is possible to continuously test using humans
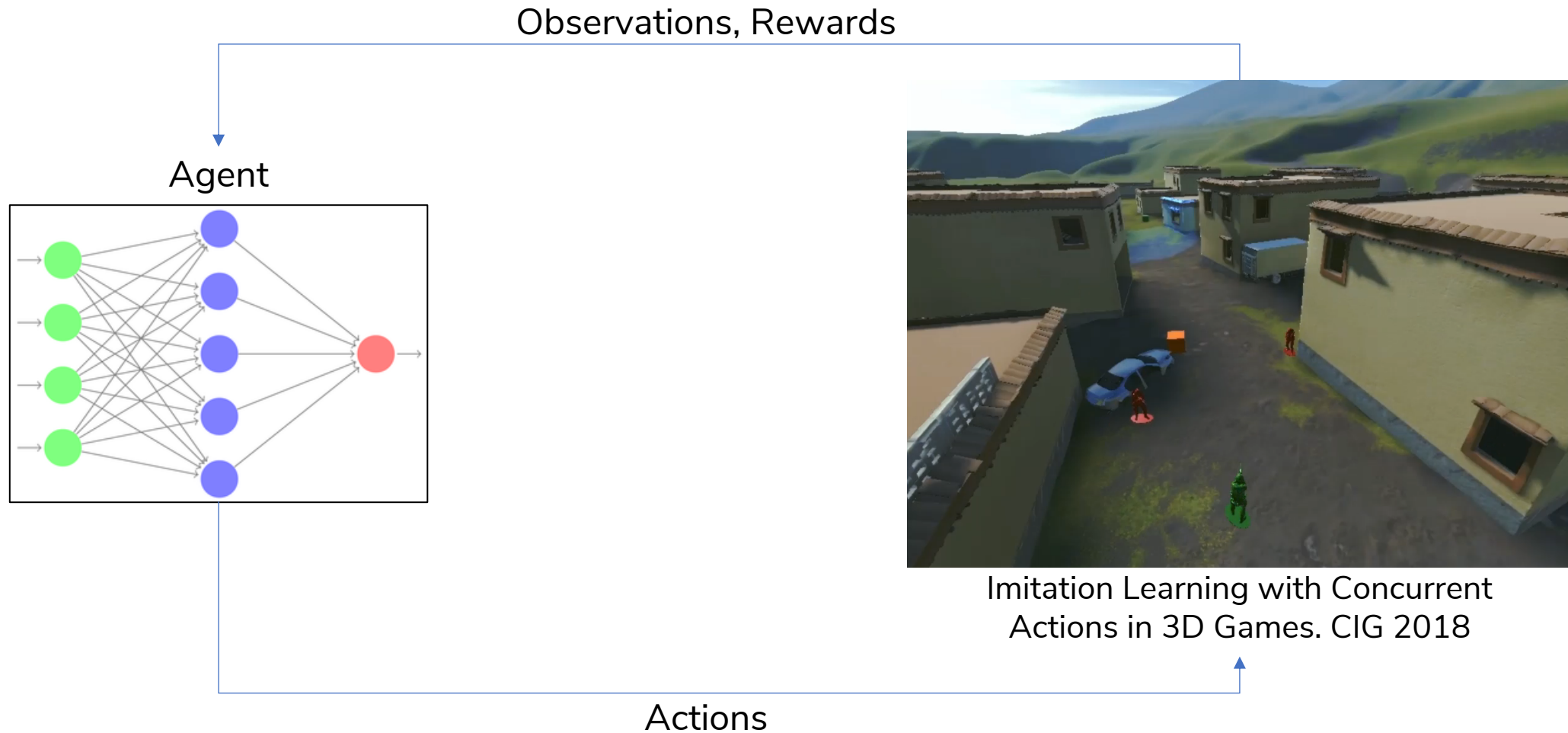
◈ Scripted AI has some drawbacks

Testing all combinations of maps and modes for 1 hour, requires 2304-man hours. That's 288 people every day!

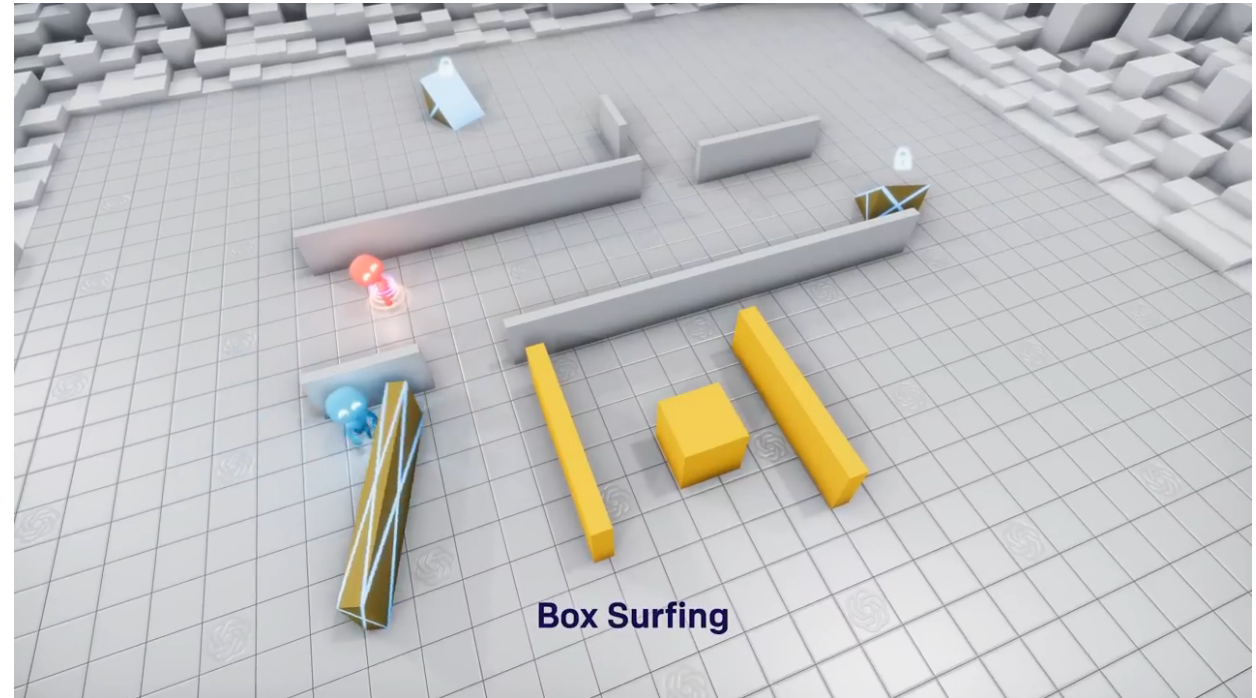*AI for Testing: The Development of Bots that Play 'Battlefield V'. Jonas Gillberg. GDC 2019*

https://www.gdcvault.com/play/1026308/AI-for-Testing-The-Development

Brief intro to RL

Observations, Rewards

Agent

Imitation Learning with Concurrent
Actions in 3D Games. CIG 2018

Actions

ADVANTAGES OF RL

	Scripted AI	RL
Adaptability	☒	✓
Trainability	☒	✓
Scalability	✓	✓
Predictability	✓	☒ 👍

OpenAI's "Hide and Seek".
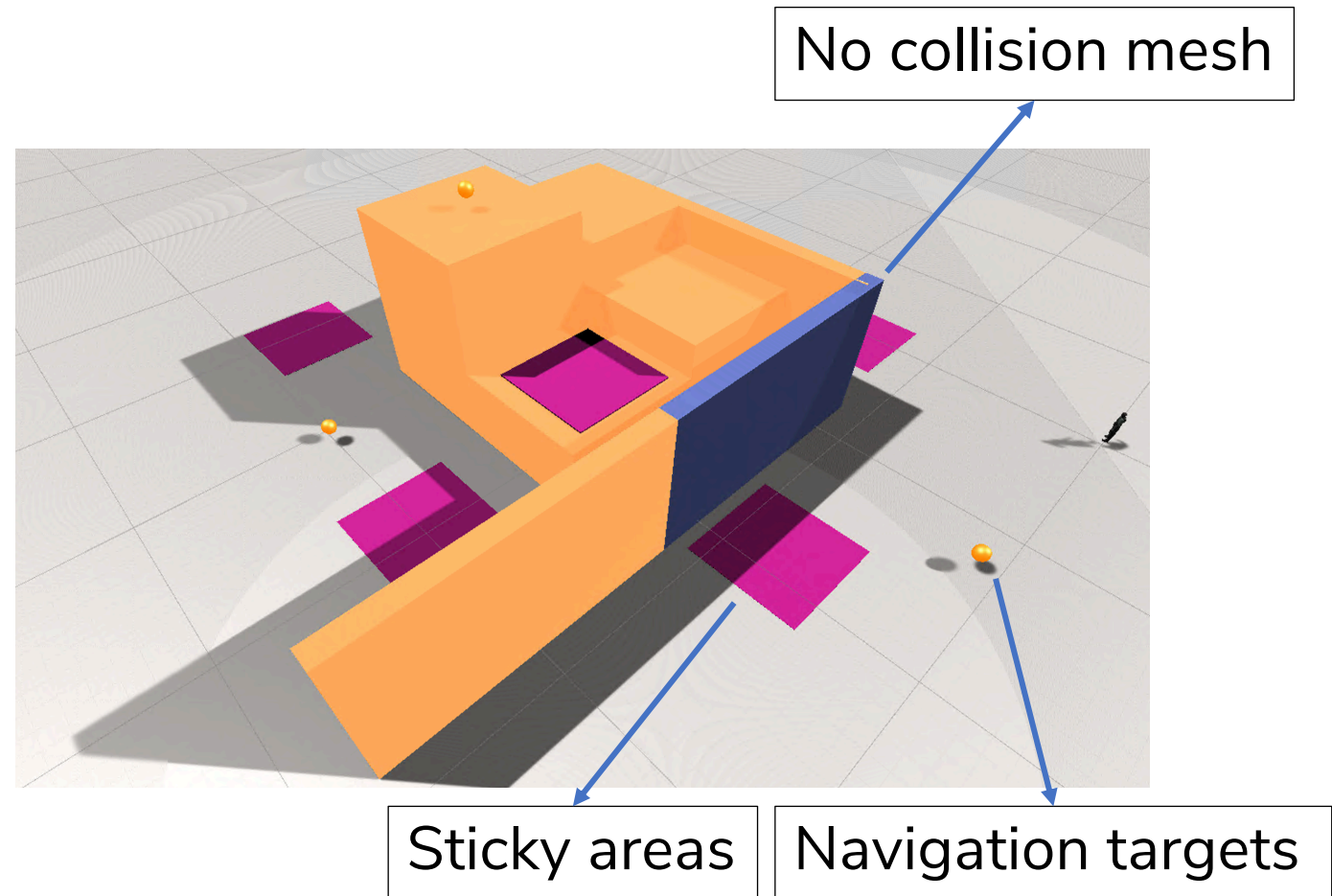Baker et al. Emergent Tool Use from Multi-Agent
Autocurricula. ICLR 2020.

ENVIRONMENTS

◈ Task
  - *Reach navigation targets*

◈ Sand-box scenarios
  - *Exploit*
  - *Stuck Player*

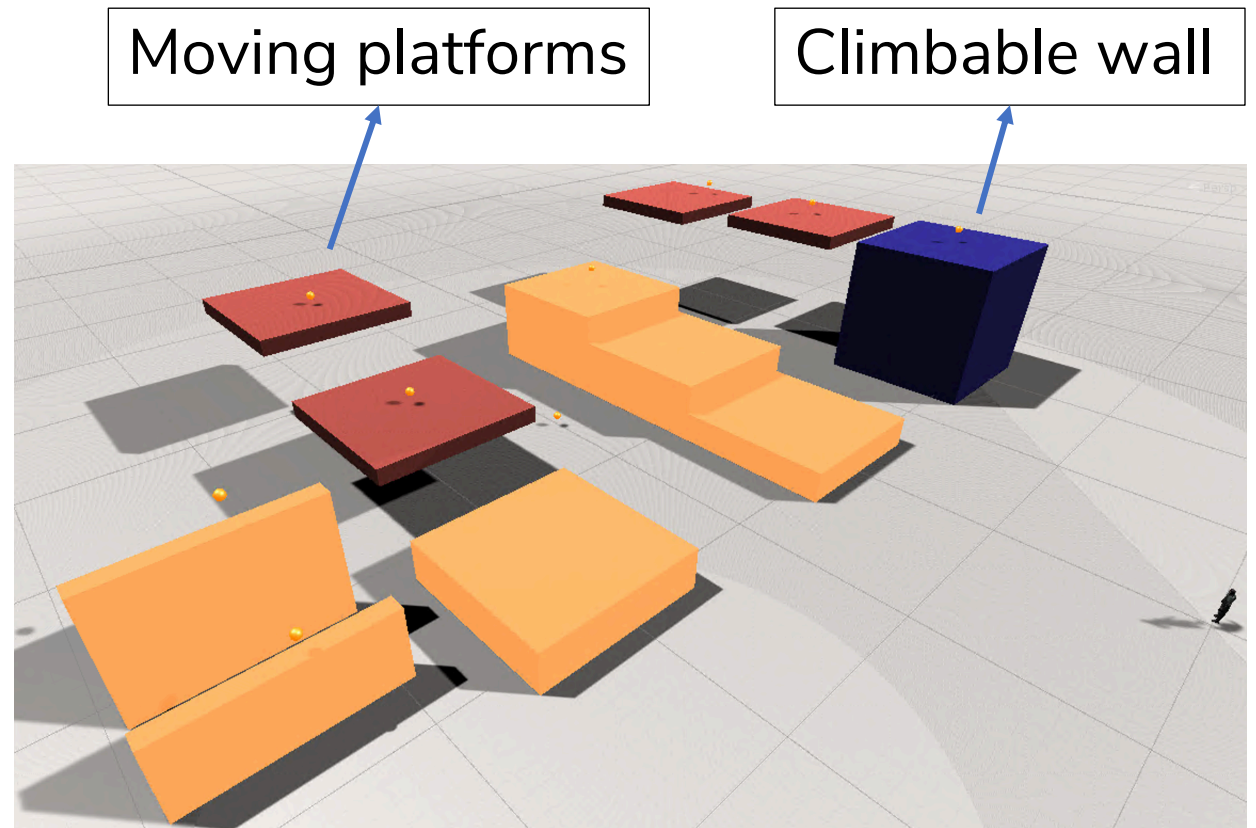 - Navigation
  - *Dynamic Navigation*

No collision mesh

Sticky areas

Navigation targets

Environments

◈ Task
- *Reach navigation targets*

◈ Sand-box scenarios
- *Exploit*
- *Stuck Player*
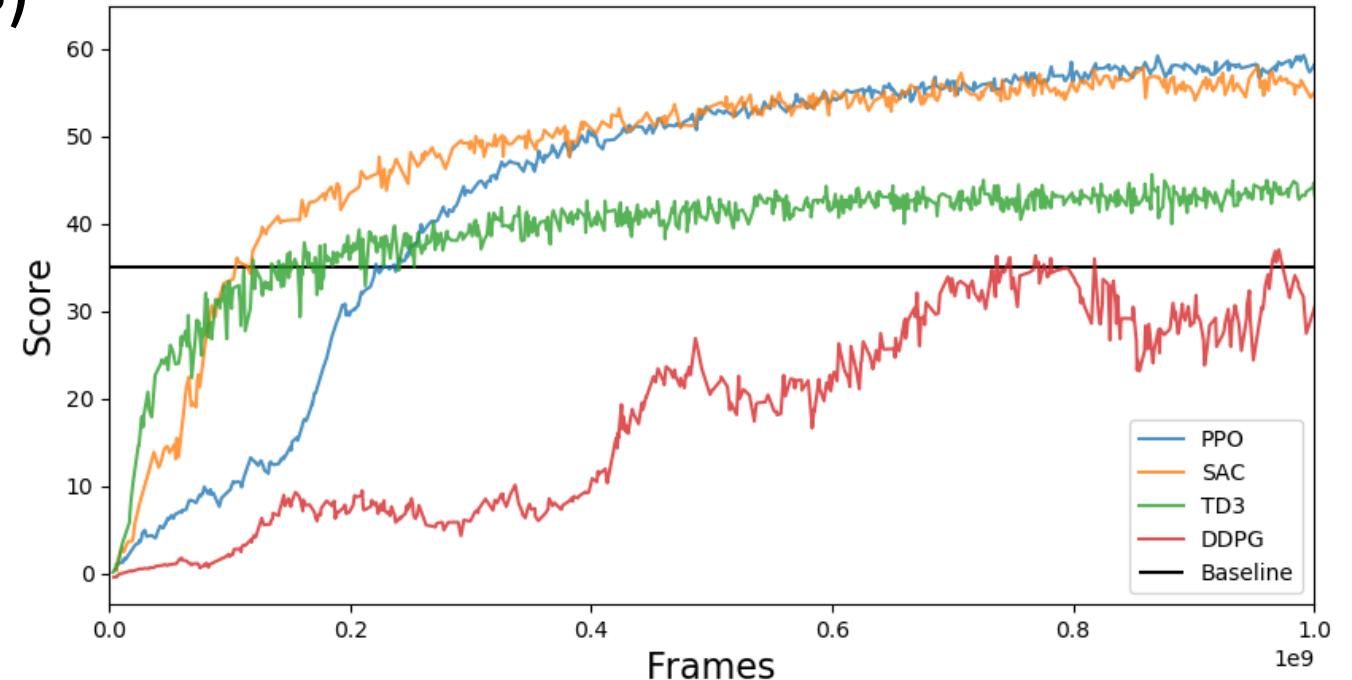
- *Navigation*
- *Dynamic Navigation*

Moving platforms

Climbable wall

ALGORITHMS

◈ Action space (continuous)
- Forward/backwards
- Left/right turning
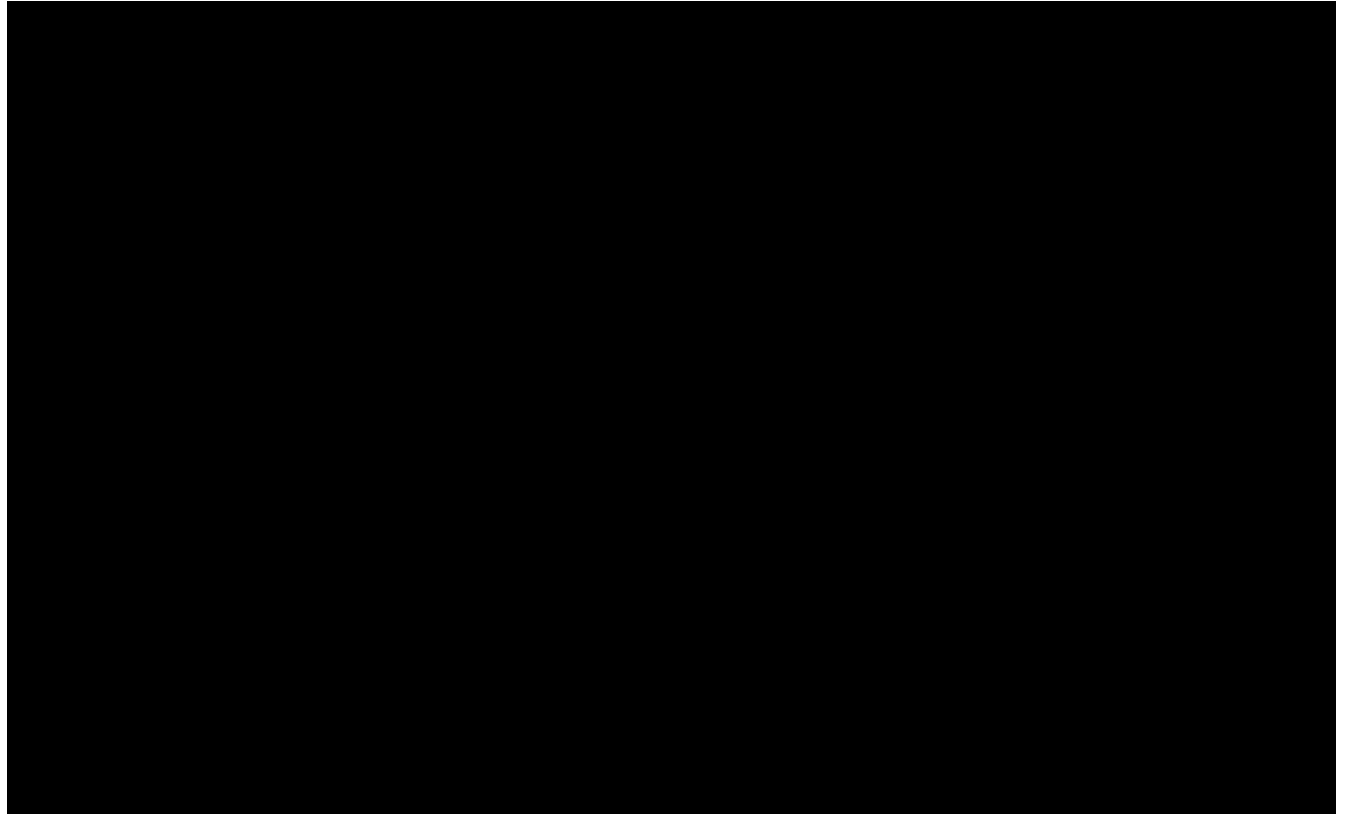- Left/right strafing
- Jump

◈ Observation
- 27 features

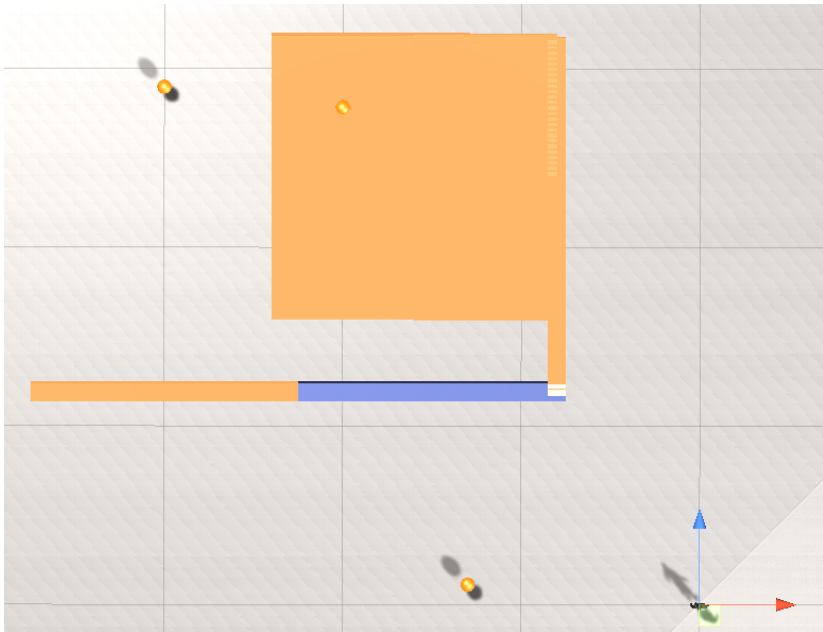◈ Reward
- $\propto \Delta_{distance\_to\_target}$

Performance when training different algorithms in the "Dynamic Navigation" sand-box

RESULTS: LARGE MAP

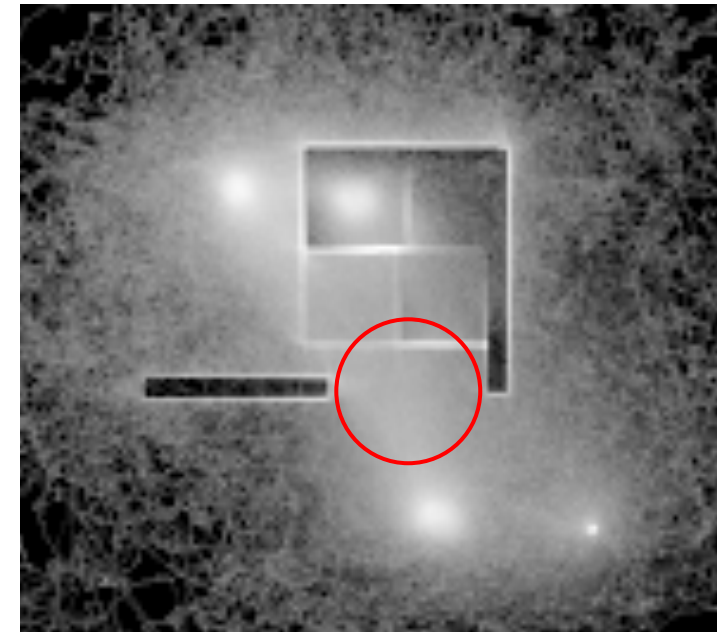| PPO	
Learning rate	$1e^{-4}$
Discount factor	0.98
PPO clip	0.2
GAE lambda	0.95
Entropy coefficient	$1e^{-2}$
V_f clip	0.2
Network architecture	Fully Connected [1024, 512, 256] + LSTM [256]

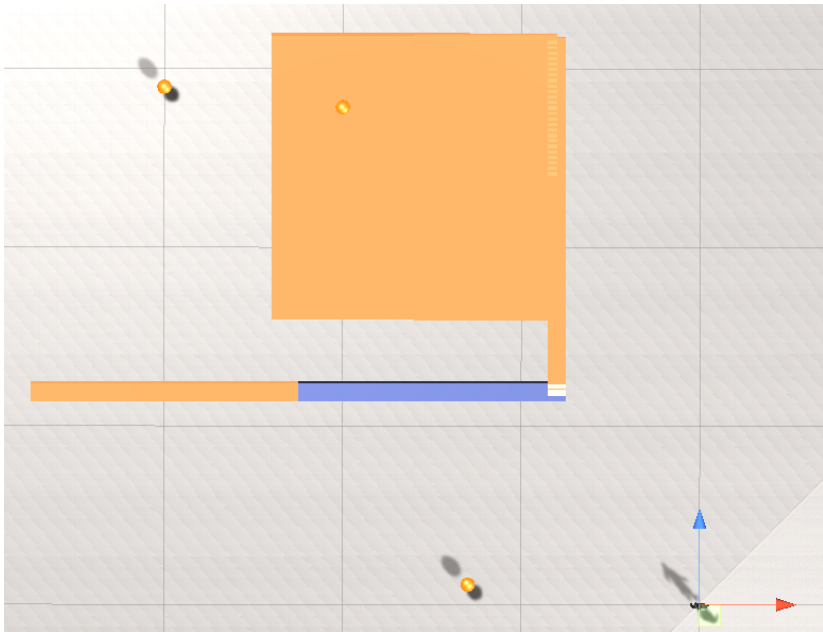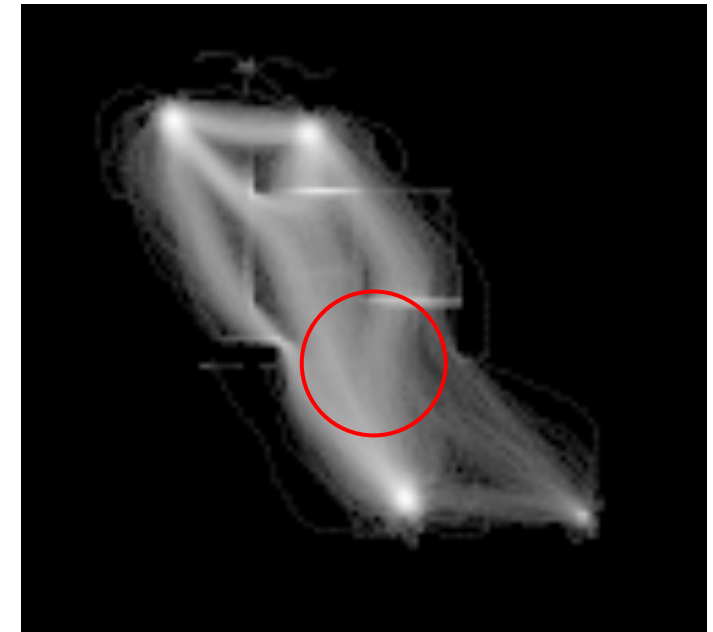# Results: "Exploit" Sand-Box

Top view of Exploit sand-box

Scripted NavMesh-based agent

RL agent after 30M steps

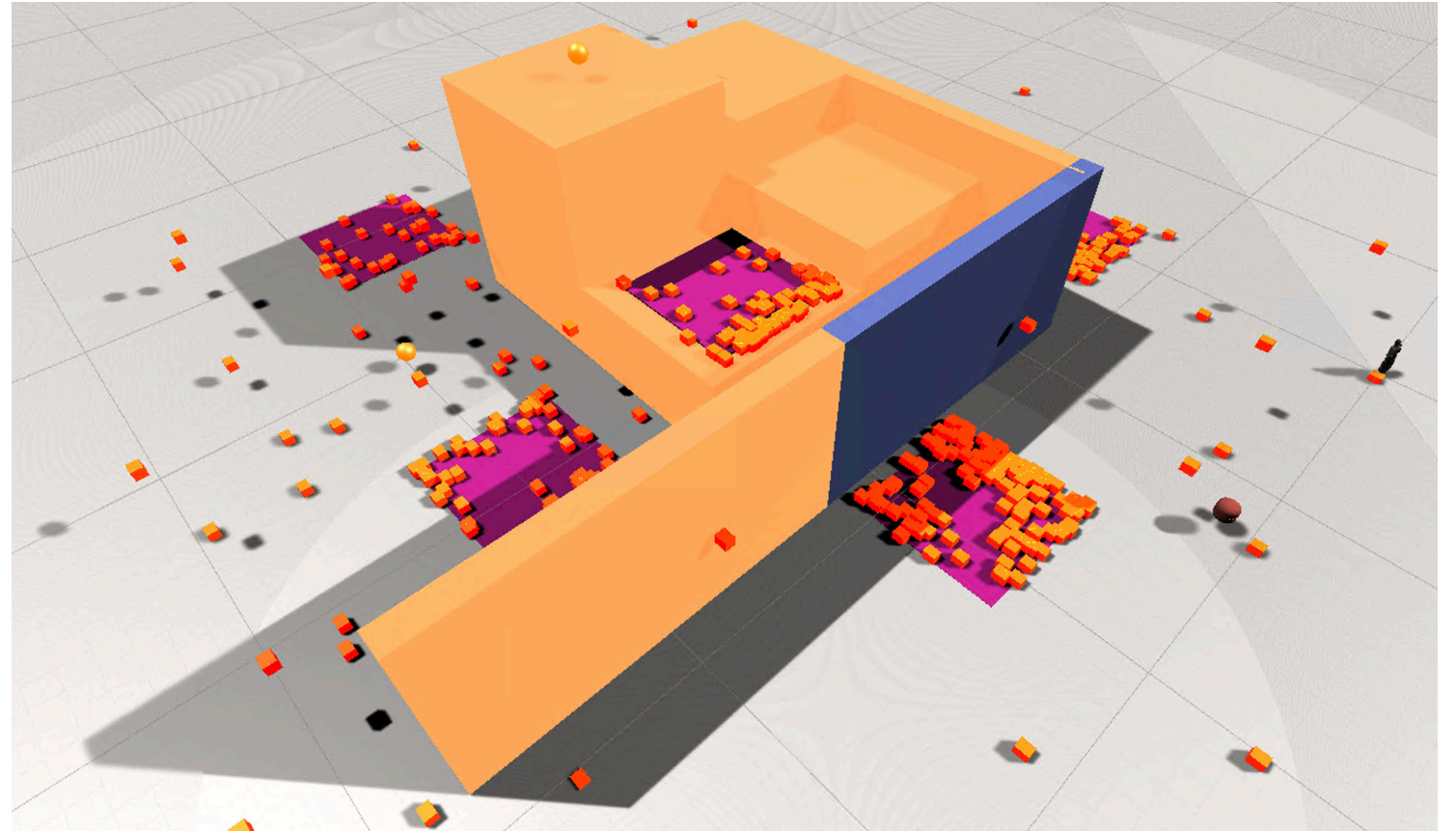# Results: "Exploit" Sand-Box
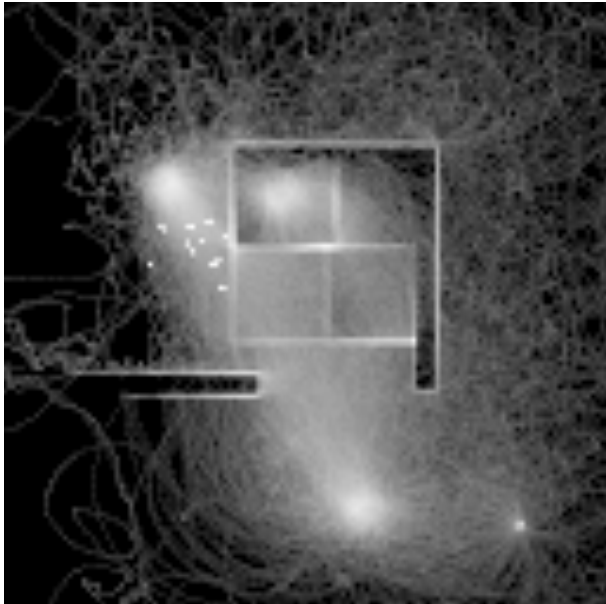

Top view of Exploit sand-box
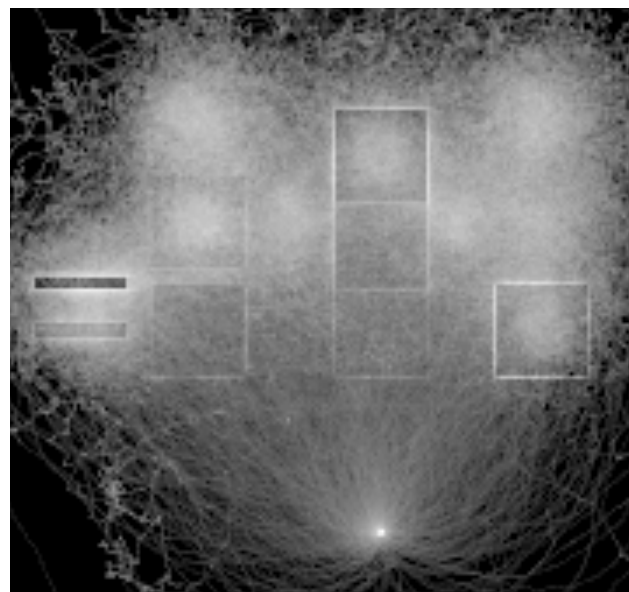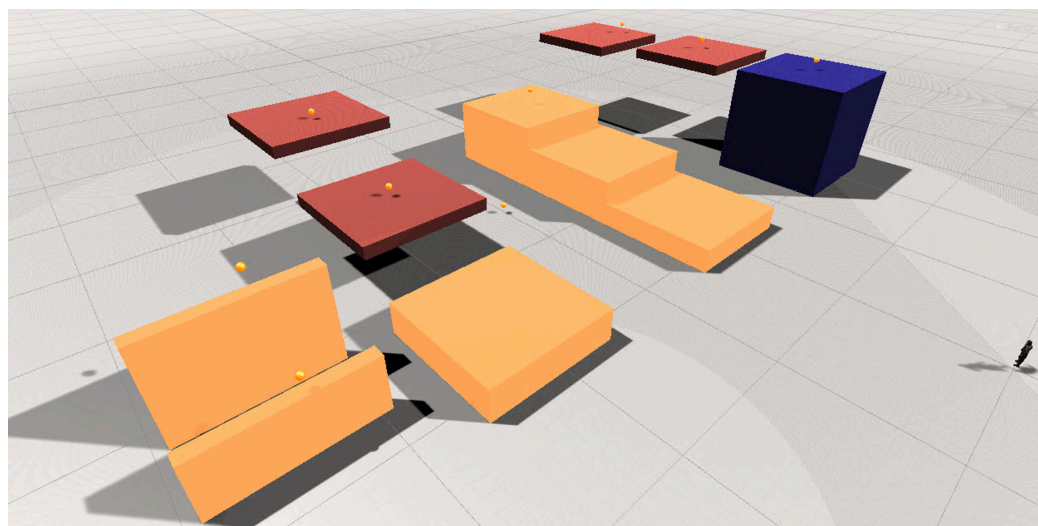
Scripted NavMesh-based agent

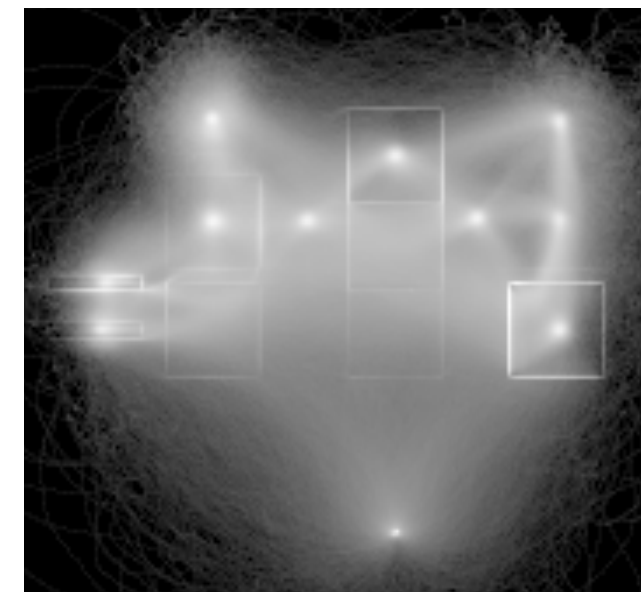RL agent fully trained

RESULTS: "STUCK PLAYER" SAND-BOX

◈ Analyzing positions when timing-out
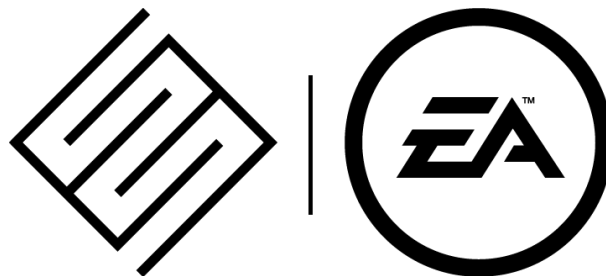
RESULTS: "DYNAMIC NAVIGATION" SAND-BOX

RL agent after 20M steps

RL agent fully trained

FUTURE WORK

◈ Data visualization and metrics

◈ Combining both scripted AI and RL

◈ Multi-Agent coordination

Joakim Bergdahl   Camilo Gordillo   Konrad Tollmar   Linus Gisslen

seed.ea.com

Stockholm – Montréal – Los Angeles – Remote

We are Hiring !