Scaling Artificial Intelligence for Digital Wargaming in Support of Decision-Making

LtCol Scotty Black, USMC

20 October 2023

“In today’s all-domain, high-tech maritime battlespace, decision advantage is arguably the difference to prevailing in war. While enabled by technology, the commander’s education, experience, and judgment become critical factors to making not only sound decisions, but to developing the cognitive capability to outthink an adversary and take decisive, bold actions—especially in combat. Intellectual overmatch is the goal. Decision advantage is the result.”

Vice Admiral Ann Rondeau, U.S. Navy (Retired)

What Do We Mean By Wargaming?

- **JP 5-0**: Representations of conflict or competition in a **synthetic environment**, in which **people make decisions** and respond to the consequences of those decisions
- **McHugh**: A **simulation**, in accordance with predetermined rules, data, and procedures, of selected aspects of a conflict situation
- **Caffrey**: wargame involves **human players or actors making decisions** in an **artificial contest environment** and then living with the consequences of their actions
- **Perla**: warfare model or **simulation** whose operation does not involve the activities of actual military forces, and whose sequence of events affects and is, in turn, affected by the **decisions made by the players** representing the opposing sides

Wargaming is about decision making in simulated combat
Wargaming provides us insight into decision-making in war

A Call for Wargaming Modernization

- The need to evolve the current paradigm of wargaming—both in terms of technology and methodology—is now being recognized throughout the Department of Defense

We must accelerate the development of AI to help us better address the complexity of modern challenges and dilemmas that currently requires human intelligence and, if possible, attempt to surpass human intelligence—***not to replace humans, but to augment and better inform human decision-making at machine speed***.

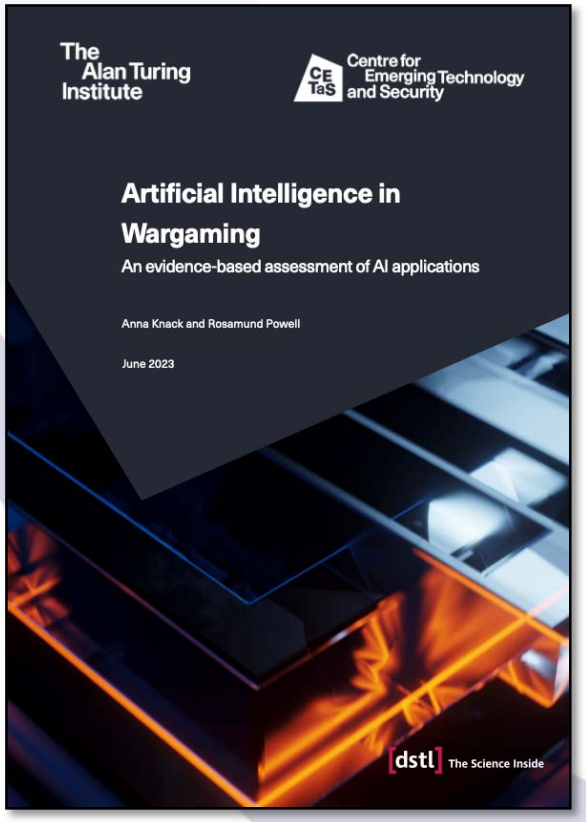
People's Republic of China Wargaming Modernization

- Our adversaries are beginning to understand the importance of wargaming
  - Pursuing innovation in platform and techniques for wargaming
  - Leveraging AI for wargaming
  - Nationwide competitions involving civilian and military

Leveraging Artificial Intelligence for Wargaming

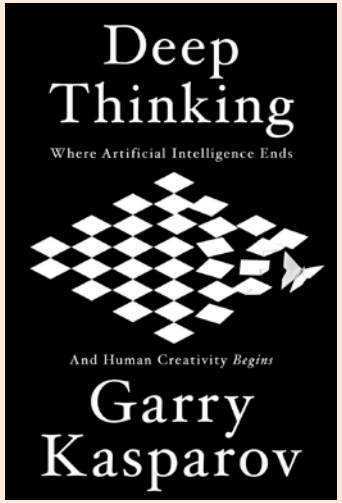
- “The rapidly improving ability of computer systems to solve problems and to perform tasks that would otherwise require human intelligence—and in some instances exceed human performance—is world altering.”
- “AI is expanding the window of vulnerability the United States has already entered.”
- “The United States must act now to field AI systems and invest substantially more resources in AI innovation to protect its security, promote prosperity, and safeguard the future of democracy.”

What Can AI Do for Wargaming?



Game design	Game execution	Game analysis	Game logistics
• AI red team agents • Procedural Content Generation • Scenario and inject generation • Gameplay testing	• Simulating opponents' responses • Adjudication • AI to run through tactical decisions • AI decision assistants and Course of Action analysis tools • Virtual assistants (to provide reminders)	• AI to highlight interesting player interactions • Data capture on player behaviour • Player decision-making analysis	• Real-time language translation • Collating past wargame data with open-source intelligence (OSINT) and secret information • Speech transcription

AI Alone is Not the Answer



After being defeated by IBM's Deep Blue, Gary Kasparov encouraged viewing AI as a tool that, when combined with human capabilities, can lead to unprecedented achievements

"*To become good at anything you have to know how to apply basic principles. To become great at it, you have to know when to violate those principles*."

- Gary Kasparov, Deep Thinking

After IBM's Watson defeated Ken Jennings and Brad Rutter (two of the best Jeopardy players ever), IBM planned on using Watson in hospitals to help diagnose and cure cancer. It failed.

"*The difference between winning at Jeopardy and curing all cancer is that we know the answer to Jeopardy questions*."

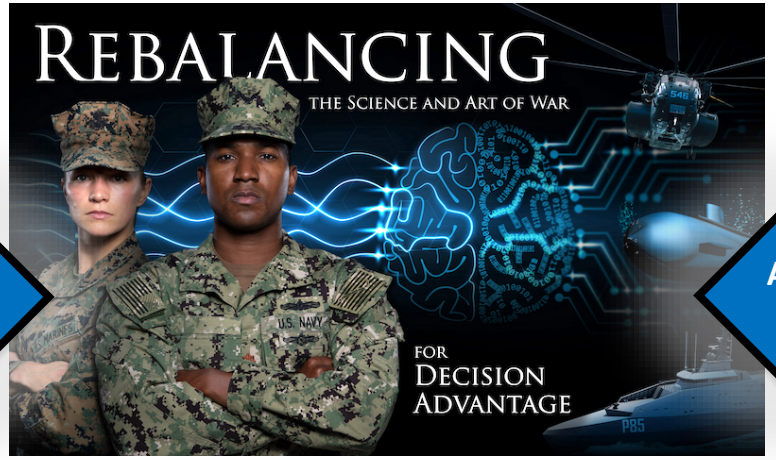
- David Epstein, Range

The Centaur Concept for Decision Making

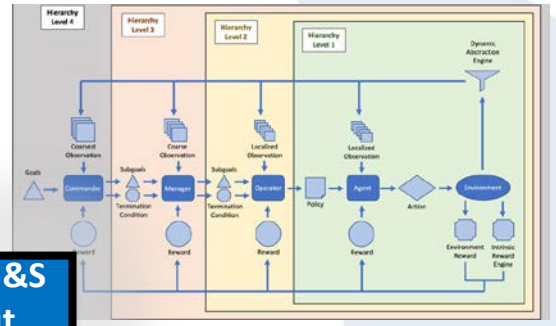
- We must begin investing heavily now in these technologies
- None of the AI gamechangers were developed overnight
  - Deep Blue took 7 years
  - Watson took 4 years
  - Game-playing AIs (AlphaGo/Zero, OpenAI Five, AlphaStar) took 2-3 years each
  - ChatGPT took 5 years

Vision

Human Input



AI + M&S Input



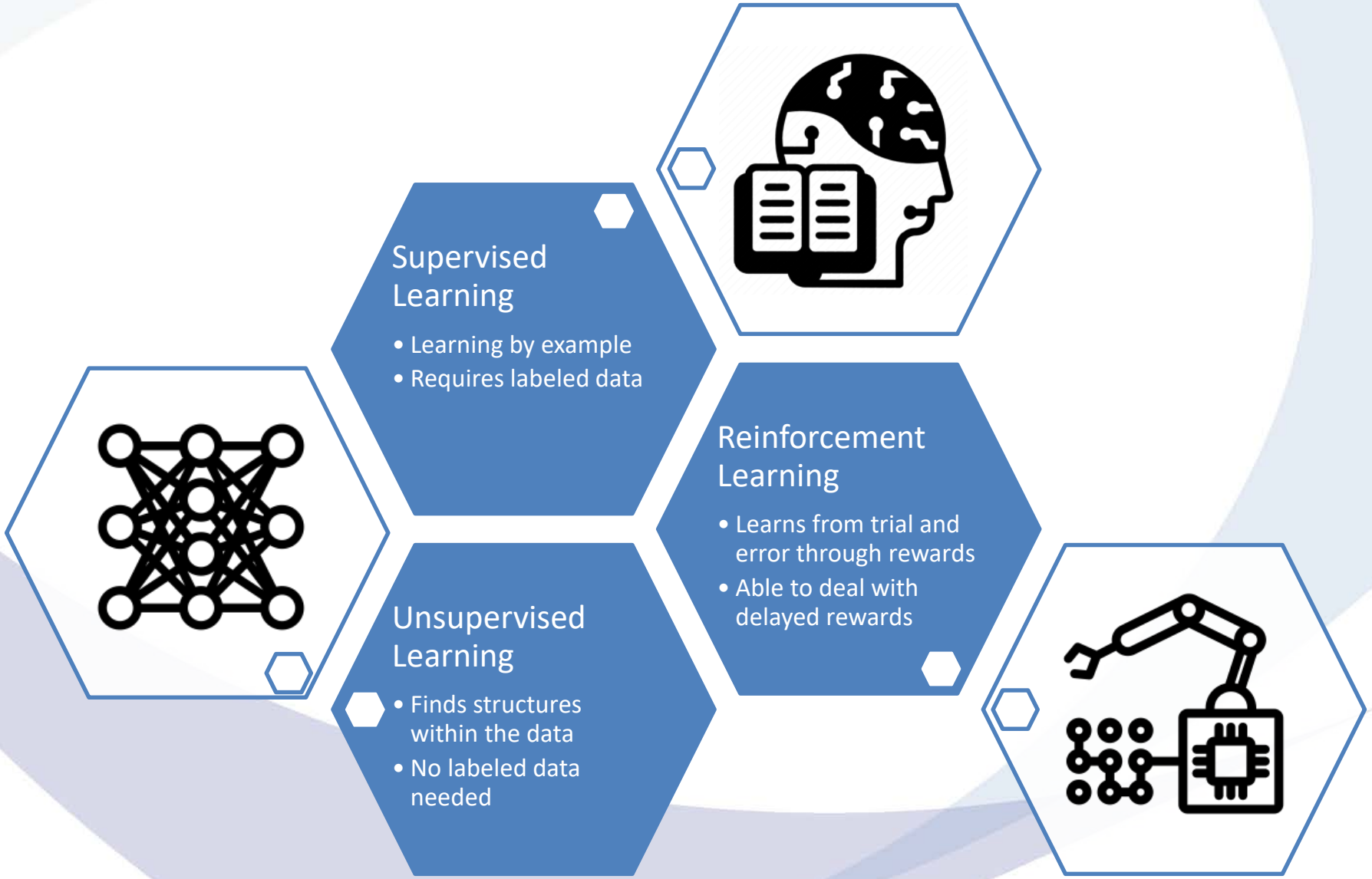
Decision Advantage

— Quality + Speed —

*Image sources: DVIDS or NPS

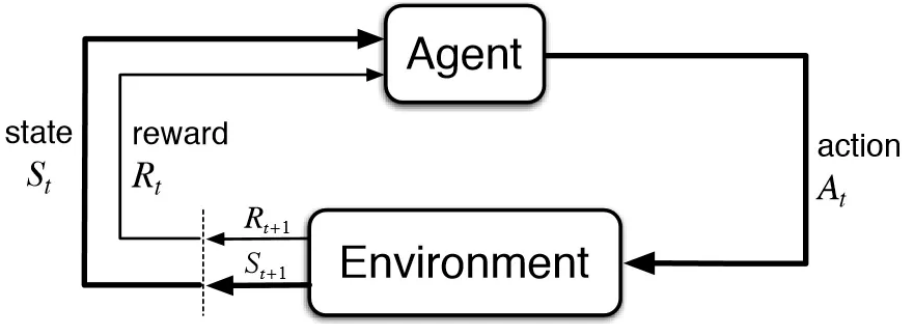
A superhuman wargaming agent serves as the foundation for creating modern decision-aid tools that can provide decision-makers more accuracy, speed, and agility over the more traditional tools

Types of Machine Learning

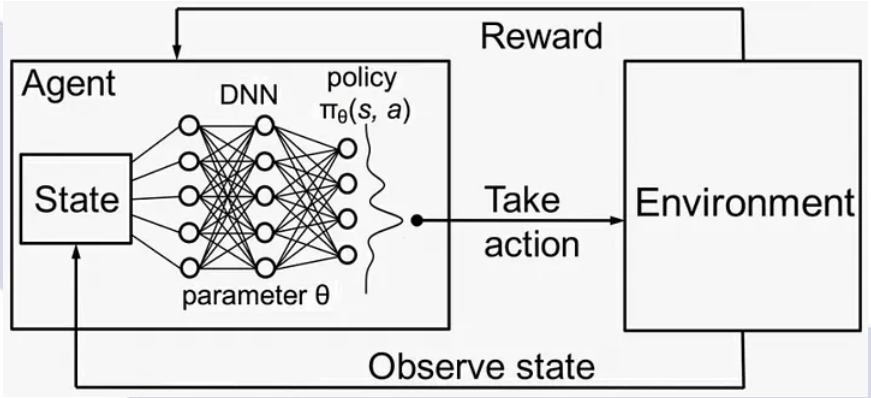


Reinforcement Learning

- Reinforcement Learning (RL)



- Deep Reinforcement Learning (DRL)

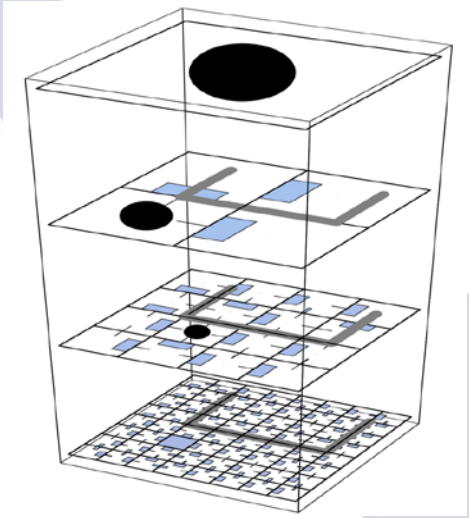


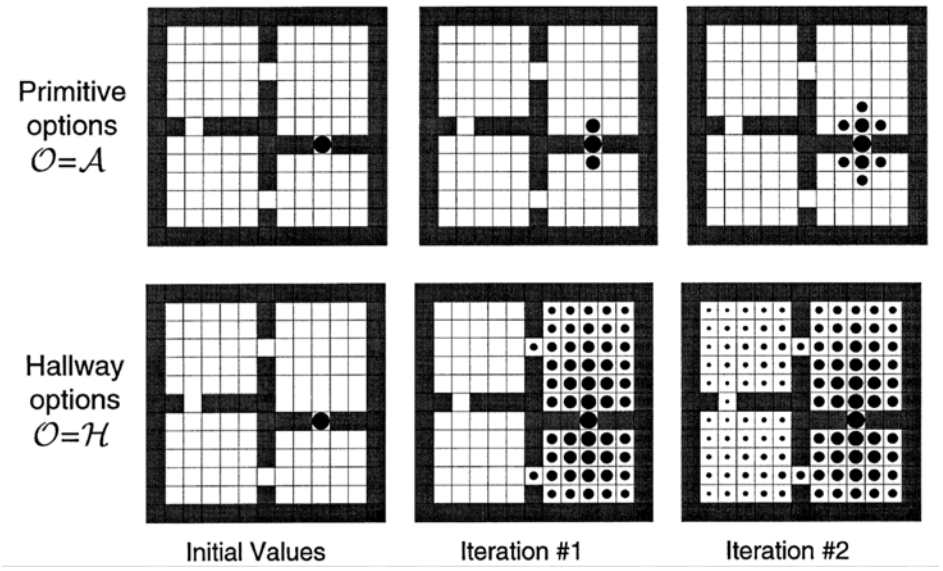
Reinforcement Learning is learning what to do through exploration and exploitation—how to map situations to actions—so as to maximize a numerical reward signal

Difference from other machine learning approaches

- Trial-and-error search
- Delayed rewards

Hierarchical Reinforcement Learning





Hierarchical Reinforcement Learning decomposes an RL problem into a hierarchy of subproblems or subtasks where a higher-level task invokes a lower-level task as if it were primitive actions

May result in solving complex problems by reducing computational complexity

Related RL Research

Commercial

Government

NPS
PRAESTANTIA PER SCIENTIAM
1909

ARMY FUTURES COMMAND
THE RESEARCH AND ANALYSIS CENTER

Naval Information Warfare Center
PACIFIC

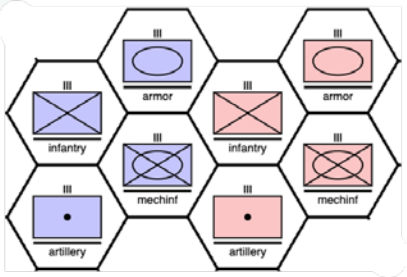
Office of Naval Research
USN
ONR
Science & Technology

DARPA

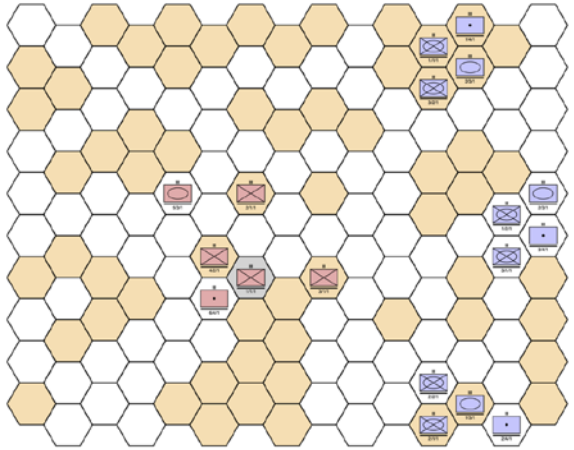
NAVSEA
WARFARE CENTERS
Indian Head

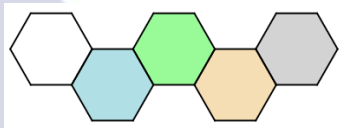
Combat Simulation Environment

- Atlatl Combat Simulation Environment

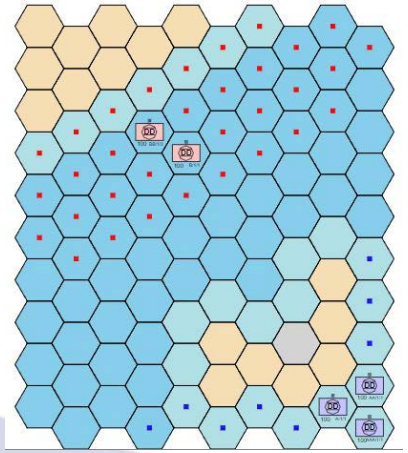


Example Unit Types

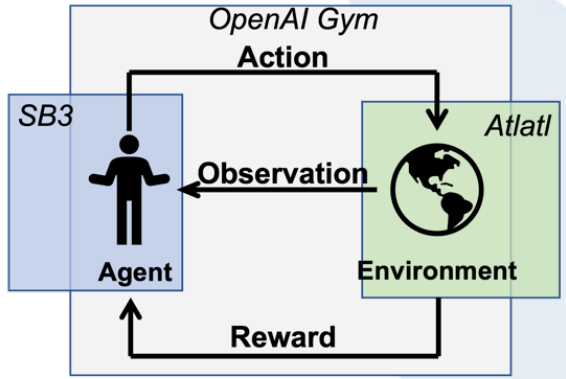




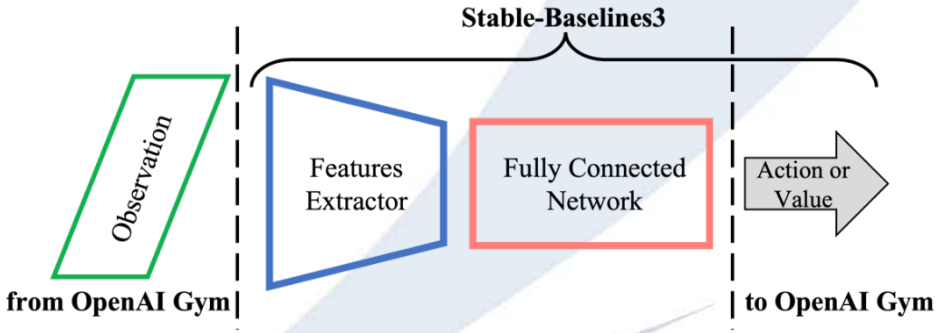
Example Terrain Types



Example Scenarios



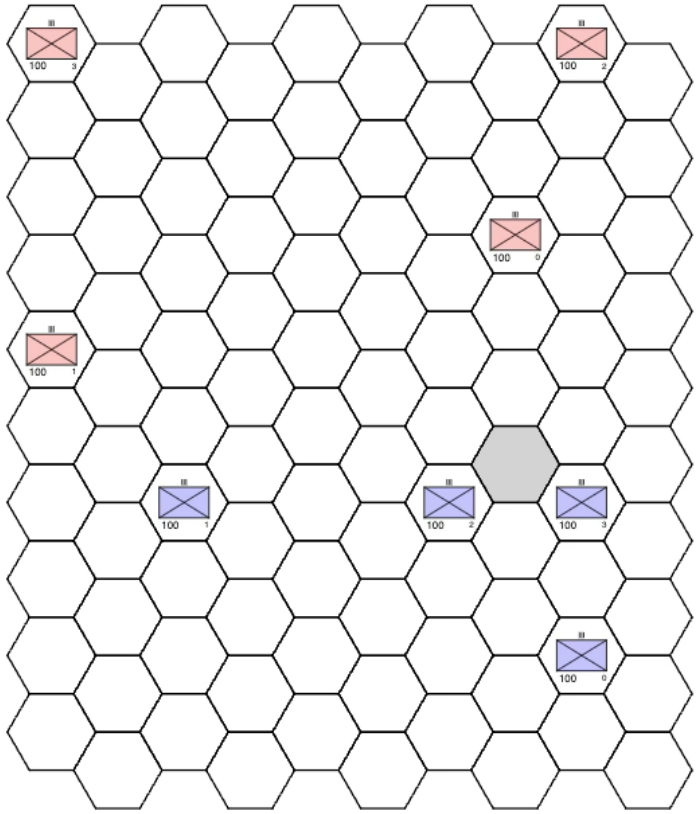
OpenAI Gym Environment (now Gymnasium)



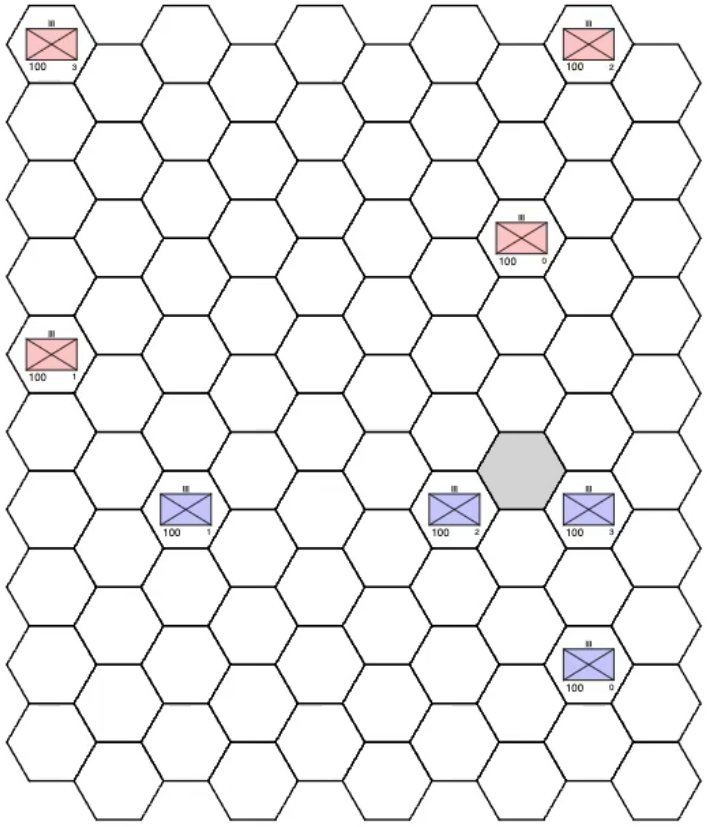
SB3 and Gym Relationship

Why RL or Machine Learning?

- Simple Example of Leveraging Machine Learning for Behavior Modeling



Scripted (Rules-Based) Agent



RL-Trained Agent

PhD Dissertation Research Areas

Research Area 1

HRL Agent Architecture and Training Framework for Combat Units

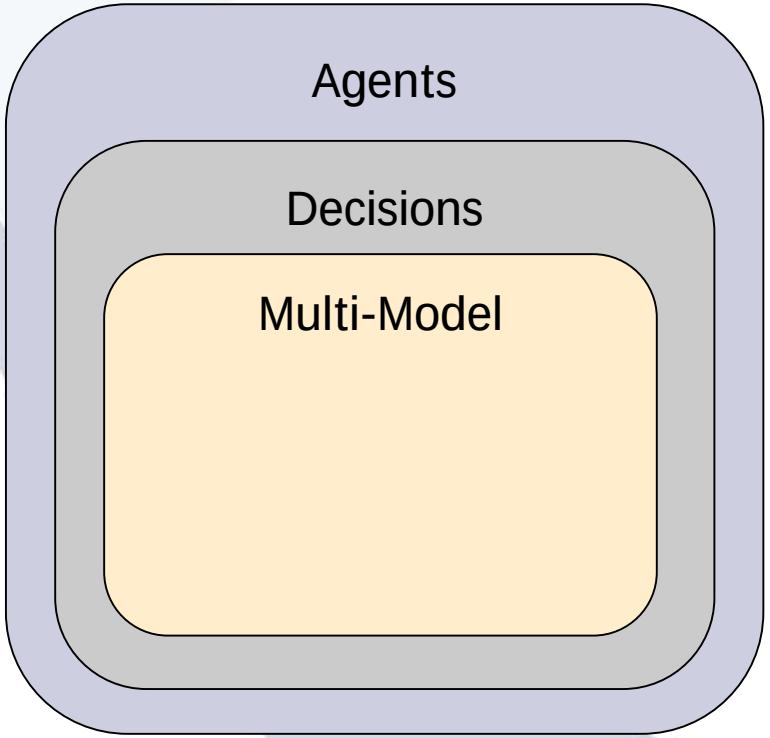
Research Area 2

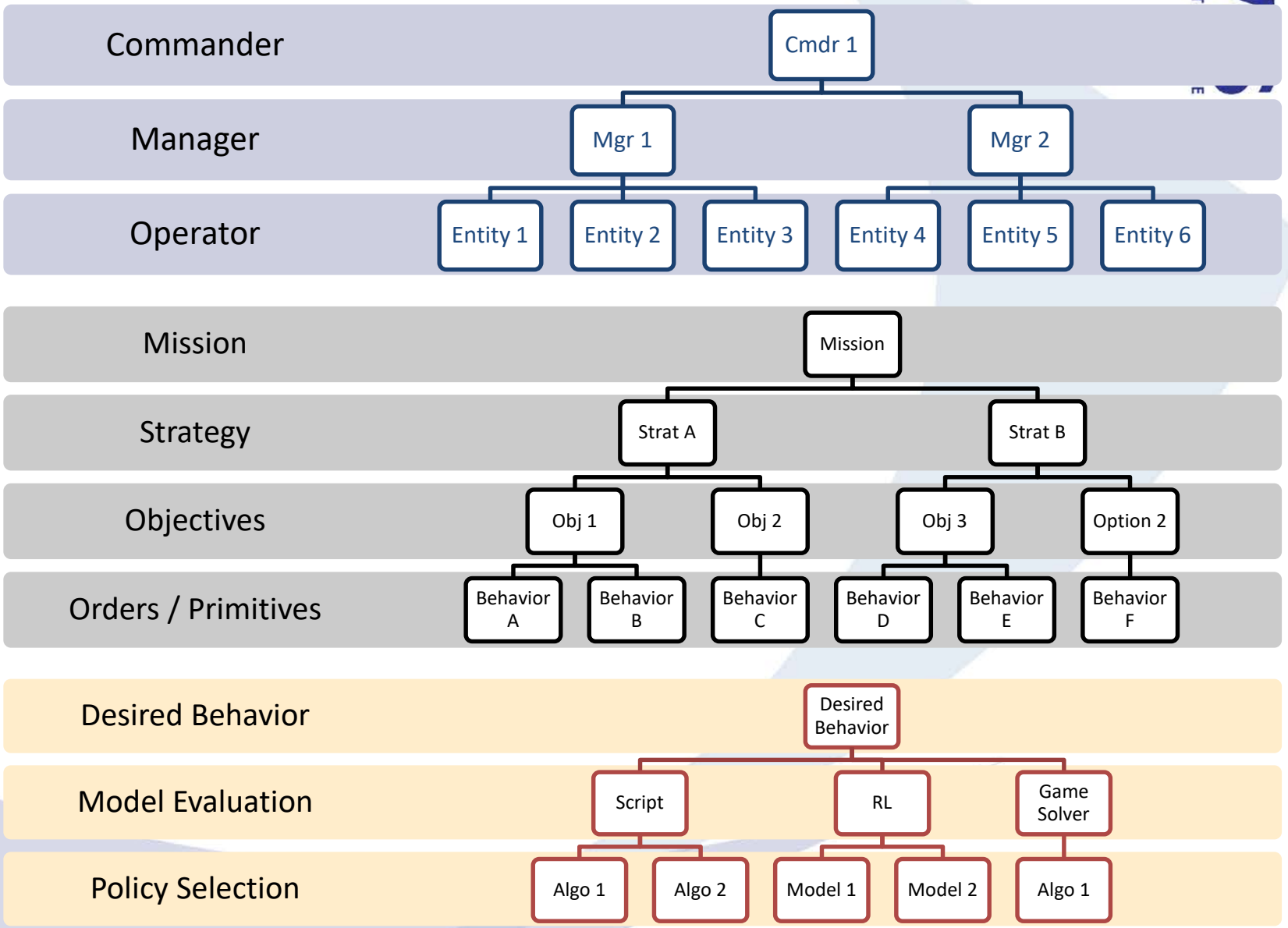
HRL Dimension-Invariant Global and Localized Observation Abstraction

Research Area 3

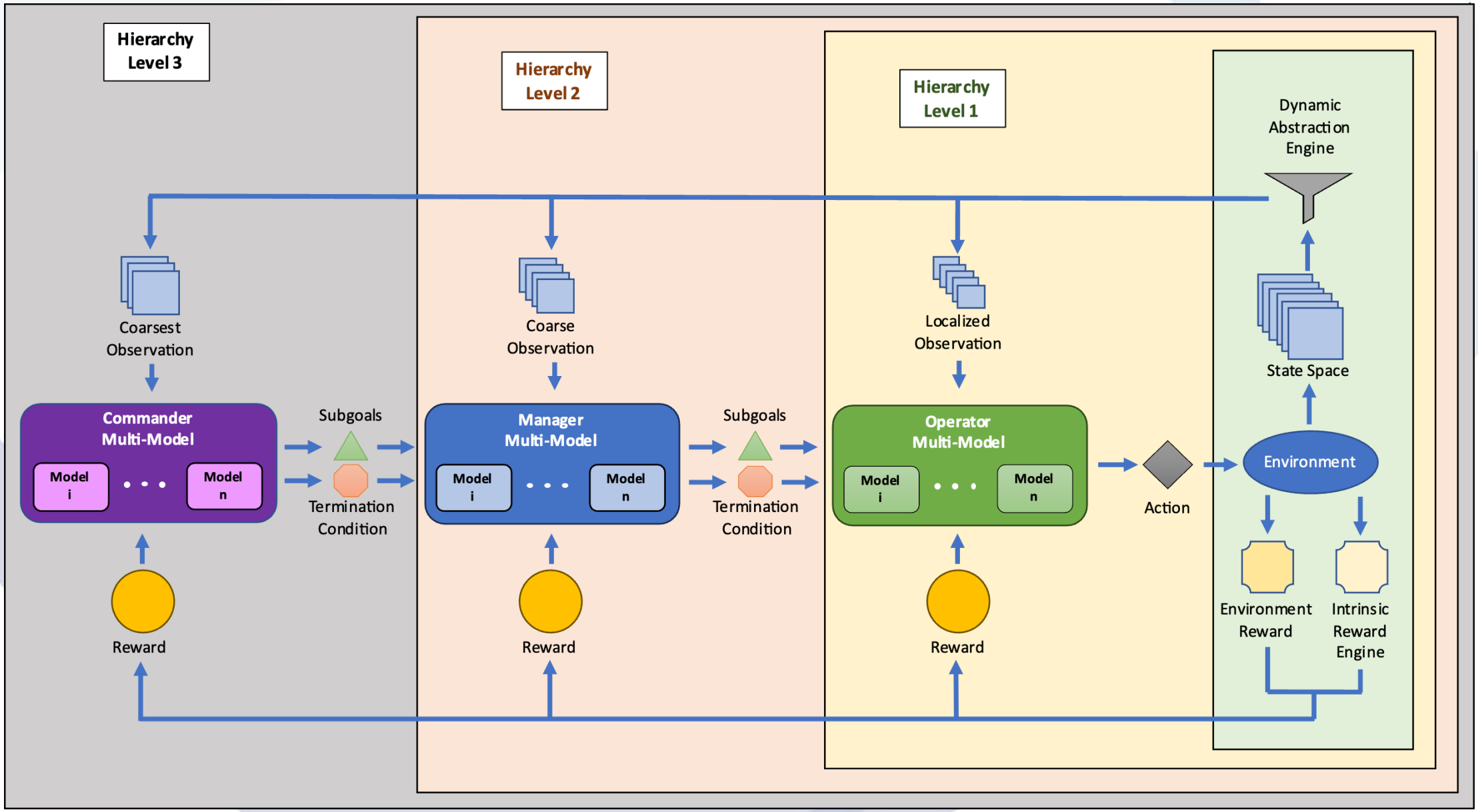
HRL Agent Architecture and Training Framework Implementation into a High-Fidelity Combat Simulation

Agent and Decision Architecture

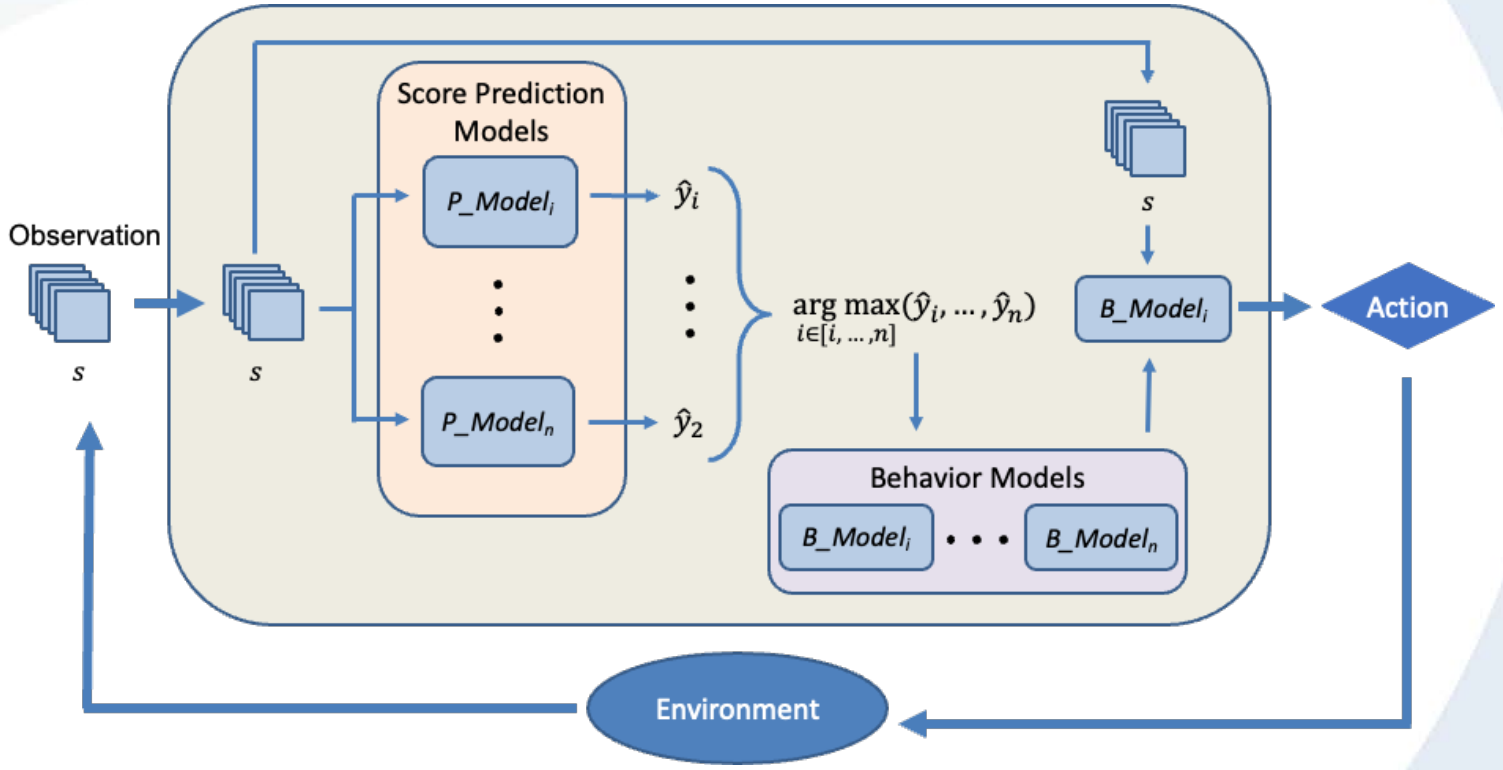




HRL Training Framework



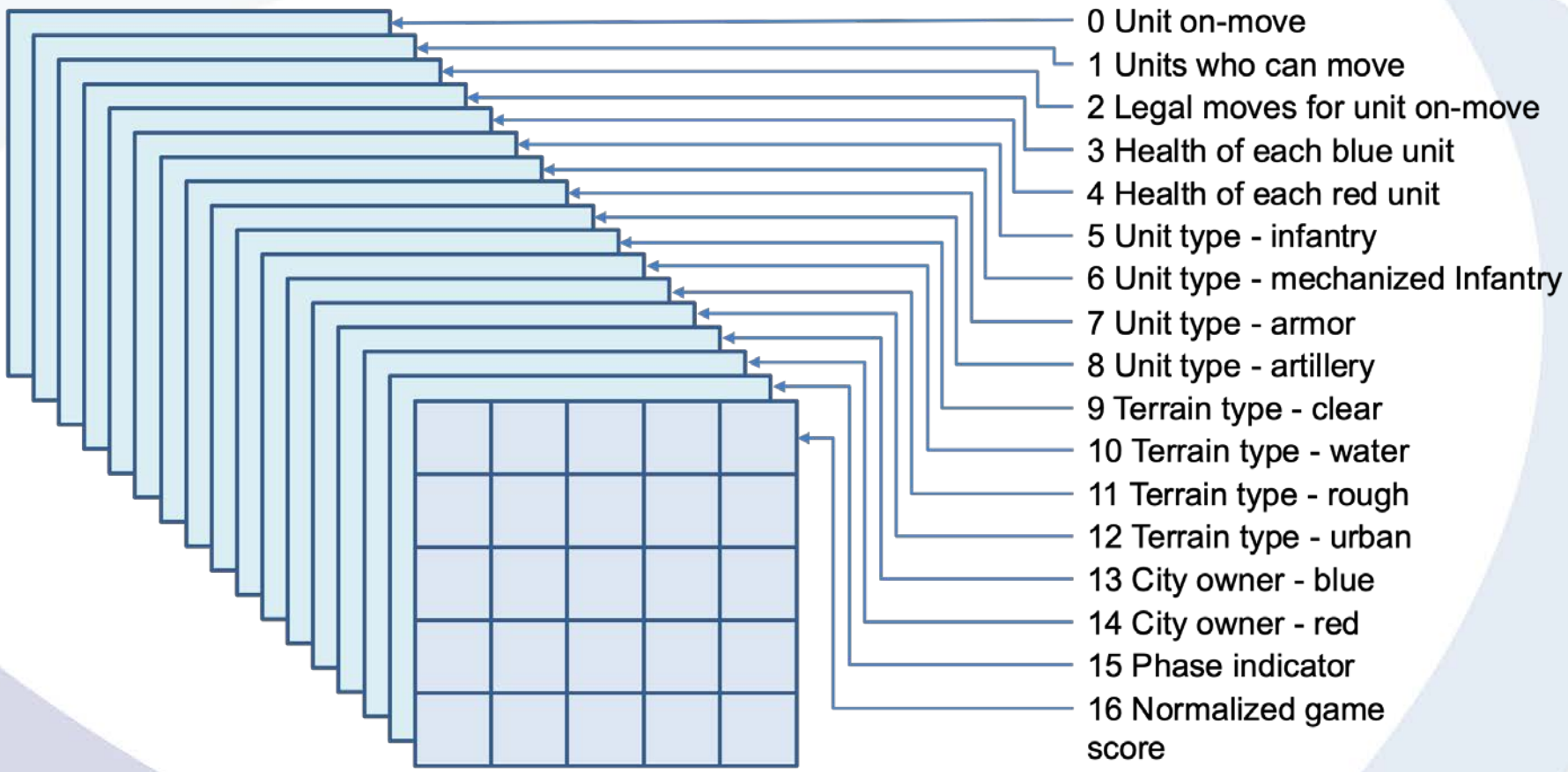
Multi-Model Framework

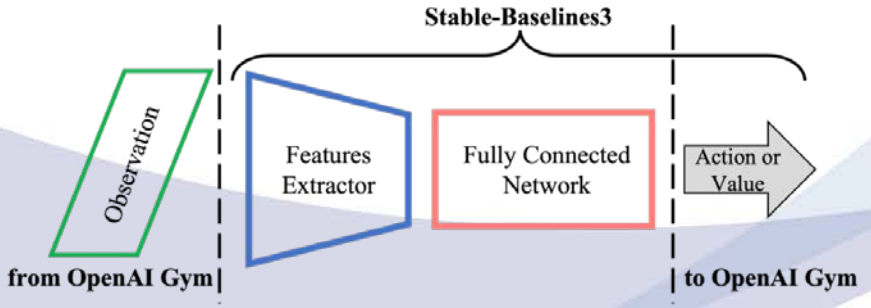


At each action-selection step, the multi-model receives an observation as input, and passes it to each of its score prediction models. Each score prediction model outputs a predicted game score, which is passed to an evaluation function. Based on this evaluation function, a behavior model is then selected. Finally, the original observation is passed to the selected behavior model and an action is output.

A multi-model approach improved the mean score by 62.6% over the mean game score of the best-performing single-model. A multi-model with more embedded behavior models outperforms a multi-model with fewer behavior models.

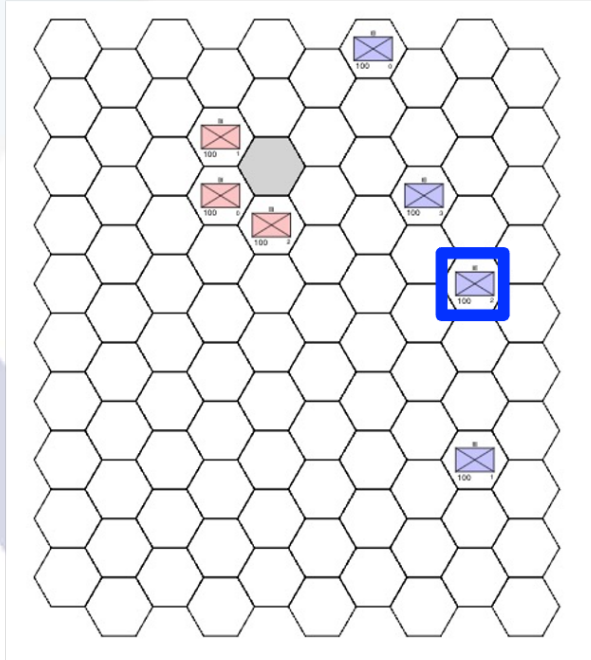
Example Observation





Abstraction Example

10x10 Scenario



*Blue box represents agent on-move

Example scenario to demonstrate concept, but scenarios can be even larger

Global State Representation

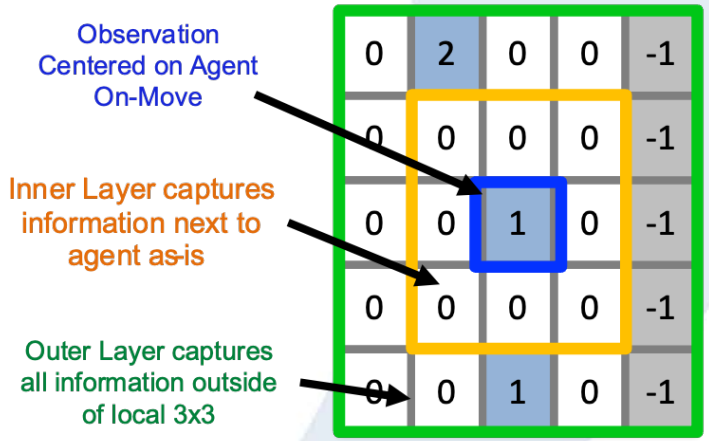
of Blue Force Channel only
(one of 17 total image channels currently)

0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

An extremely hard problem to solve as-is, and can quickly become an intractable problem for bigger, more realistic scenarios

Local 5x5 State Representation

of Blue Force Channel only
(one of 17 total image channels currently)

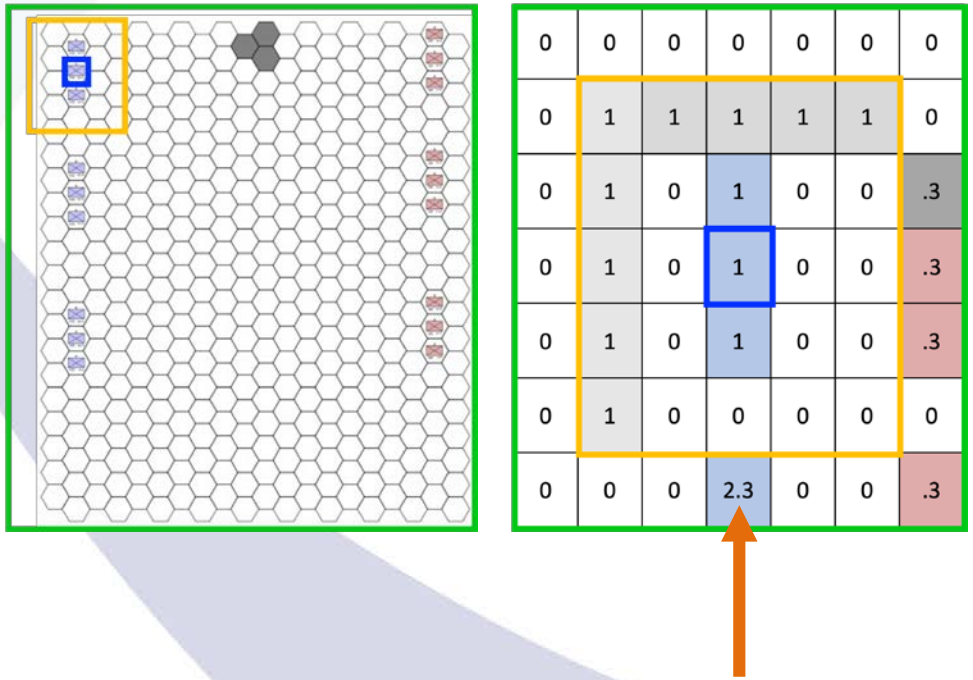


*** Also experimenting with 7x7 and other dimensions

Can potentially make any problem more tractable

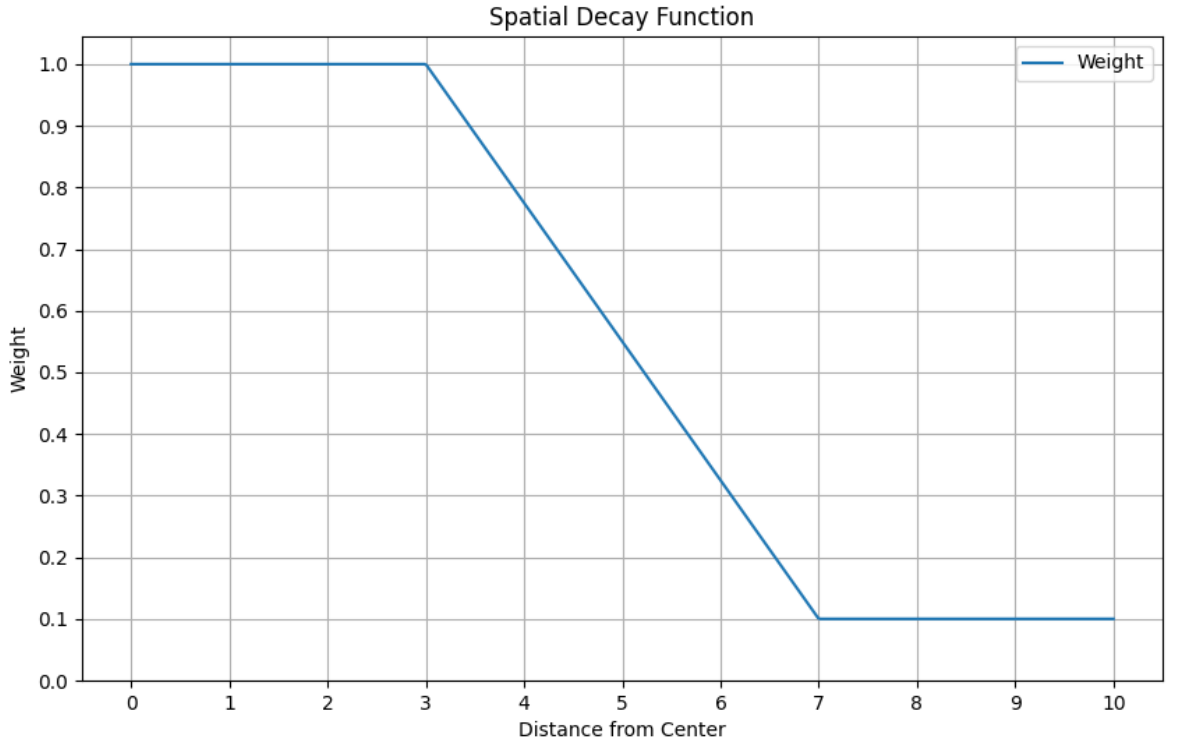
Localized Observation

- Abstraction via Local Piecewise Linear Spatial Decay



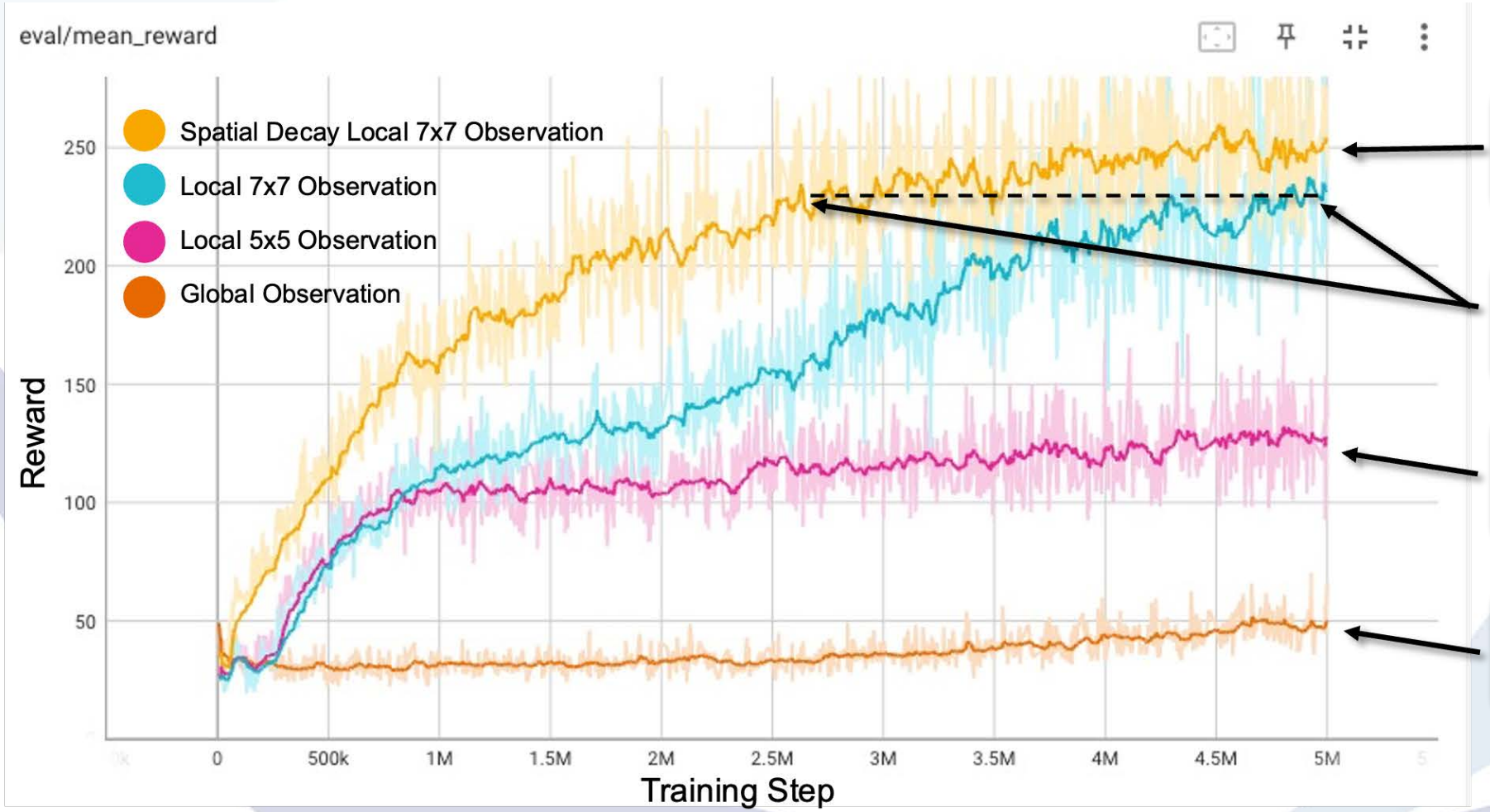
Values in the outermost layer now use weights computed via a piecewise linear spatial decay.

These weights are then multiplied by the values of the respective layers and then summed per radial.



Piecewise linear spatial decay function

Learning Rate Comparison



Using localized 7x7 with spatial decay resulted in most learning

Using spatial decay resulted in learning the same in about half the time steps

5x5 observation plateaued early-on in training

Full global observation of 10x10 space resulted in very slow learning

Conclusion

- We must accelerate research in this domain
- The AI is not meant to replace humans, but augment and inform human decision-making
- Intelligent agents in simulation are key to developing decision-aid tools
- Reinforcement learning is just one approach, others should also be investigated

Questions / Comments / Discussion

LtCol Scotty Black, USMC
scotty.black@nps.edu

Chris Darken, PhD
cjdarken@nps.edu