Current State and Future of AI:
Separating Fact from Fiction

Kagan Tumer

Director, Collaborative Robotics and Intelligent Systems (CoRIS) Institute

Professor, Mechanical Engineering & Computer Science

Oregon State University

kagan.tumer@oregonstate.edu

July 15, 2018

Oregon State University

Disclaimer … Plea

All examples of AI systems, companies, and movies are meant to showcase capabilities and discuss directions for improvement

Comments not meant to disparage or criticize

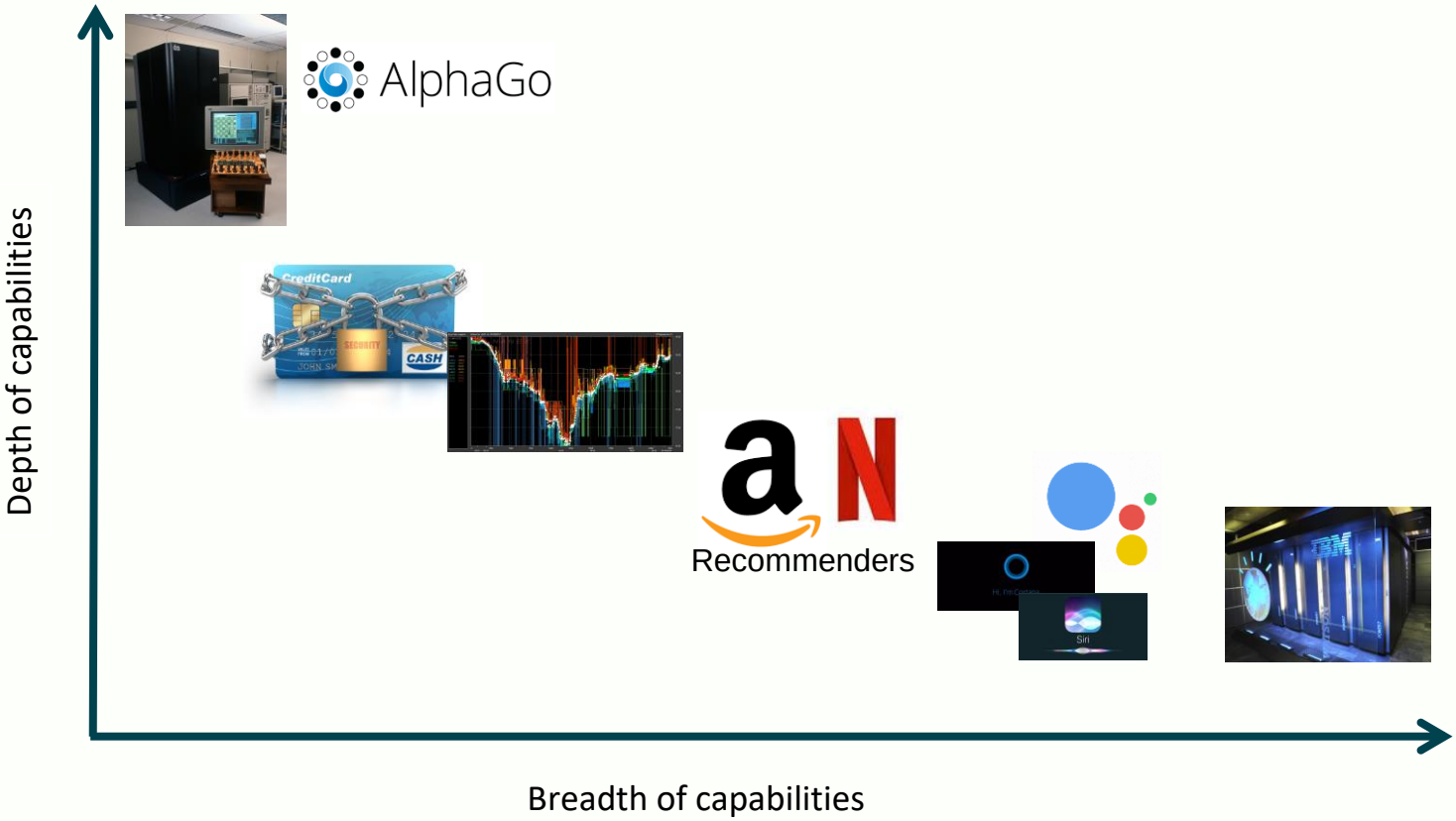
As a community, can we please stop saying:

“but that’s not real AI”

every time we have a success story?

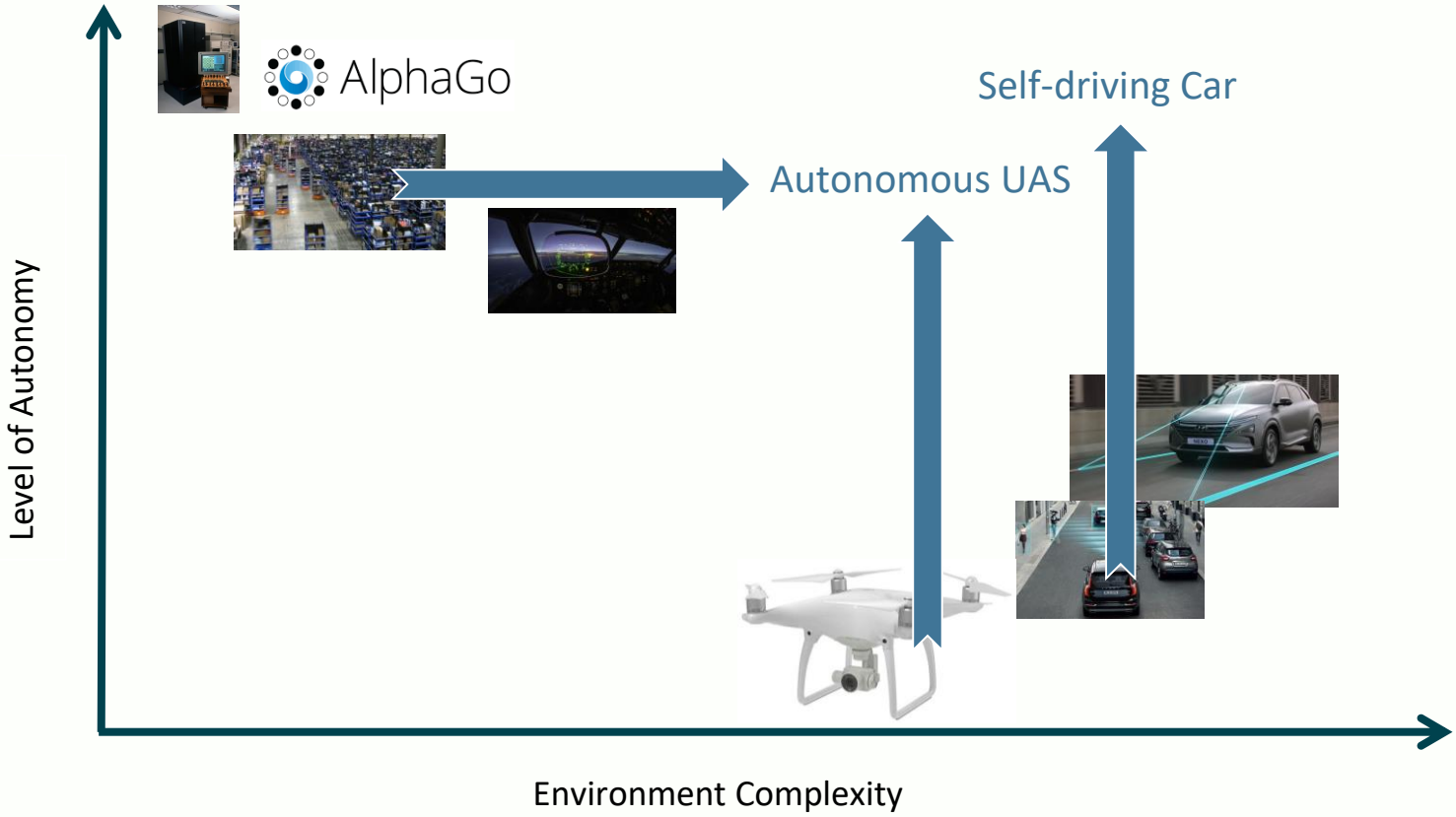
Oregon State University

Where are we ?



Oregon State University

Where are we ?



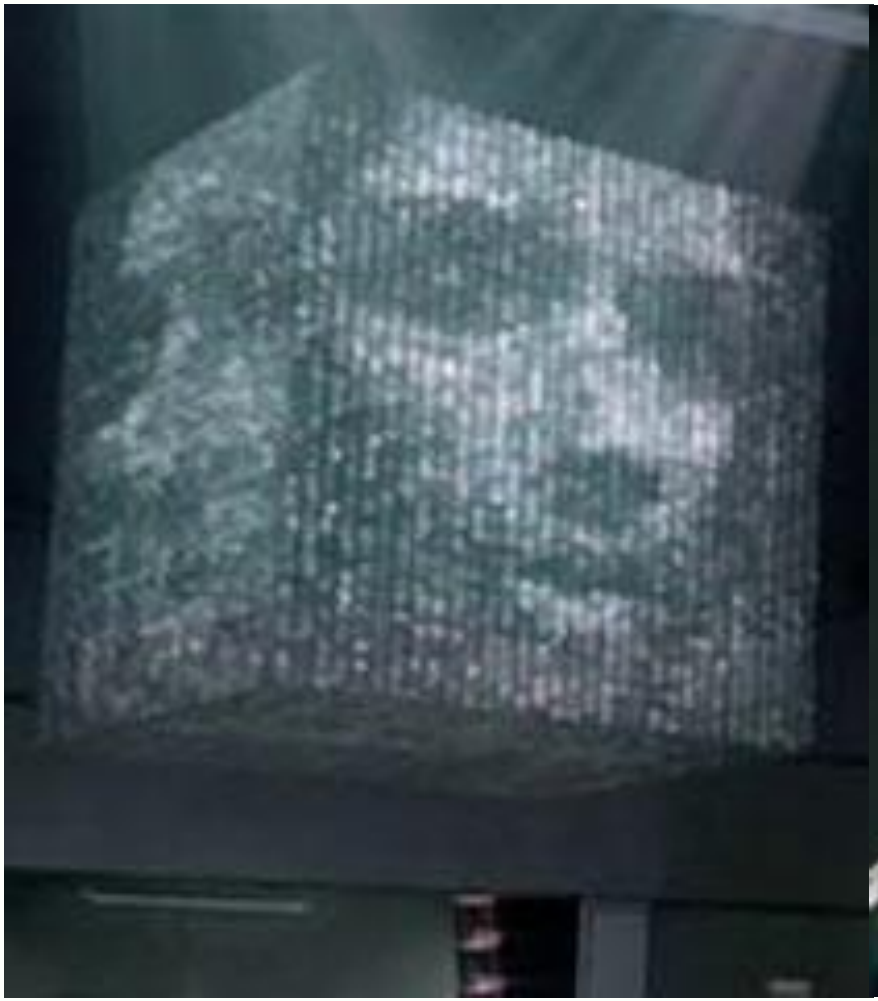
Oregon State University

Where are we ?

Where do we want to go?

Oregon State University

How is where we're going depicted ?



Oregon State University

What will it take to get where we're going ?
(and how to NOT go where we don't want to go)

Oregon State University

What will it take to get where we're going ?

Reward Shaping: D++

Curriculum Learning

Memory Block

Multi-Reward Learning

Long-term autonomy

Few-shot learning

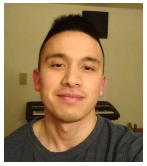
Broad AI

High competency

Complex Environment

Explanations

Suggestions



Oregon State University

Four Directions

How do I complement my teammates?

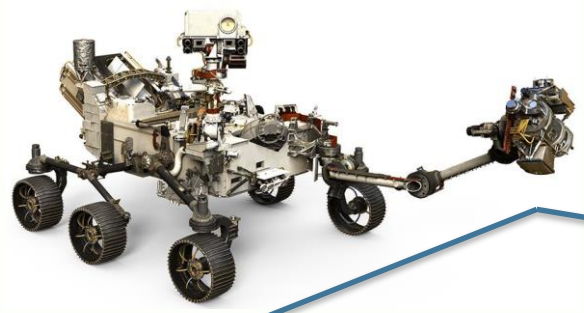
Walk before I run ?

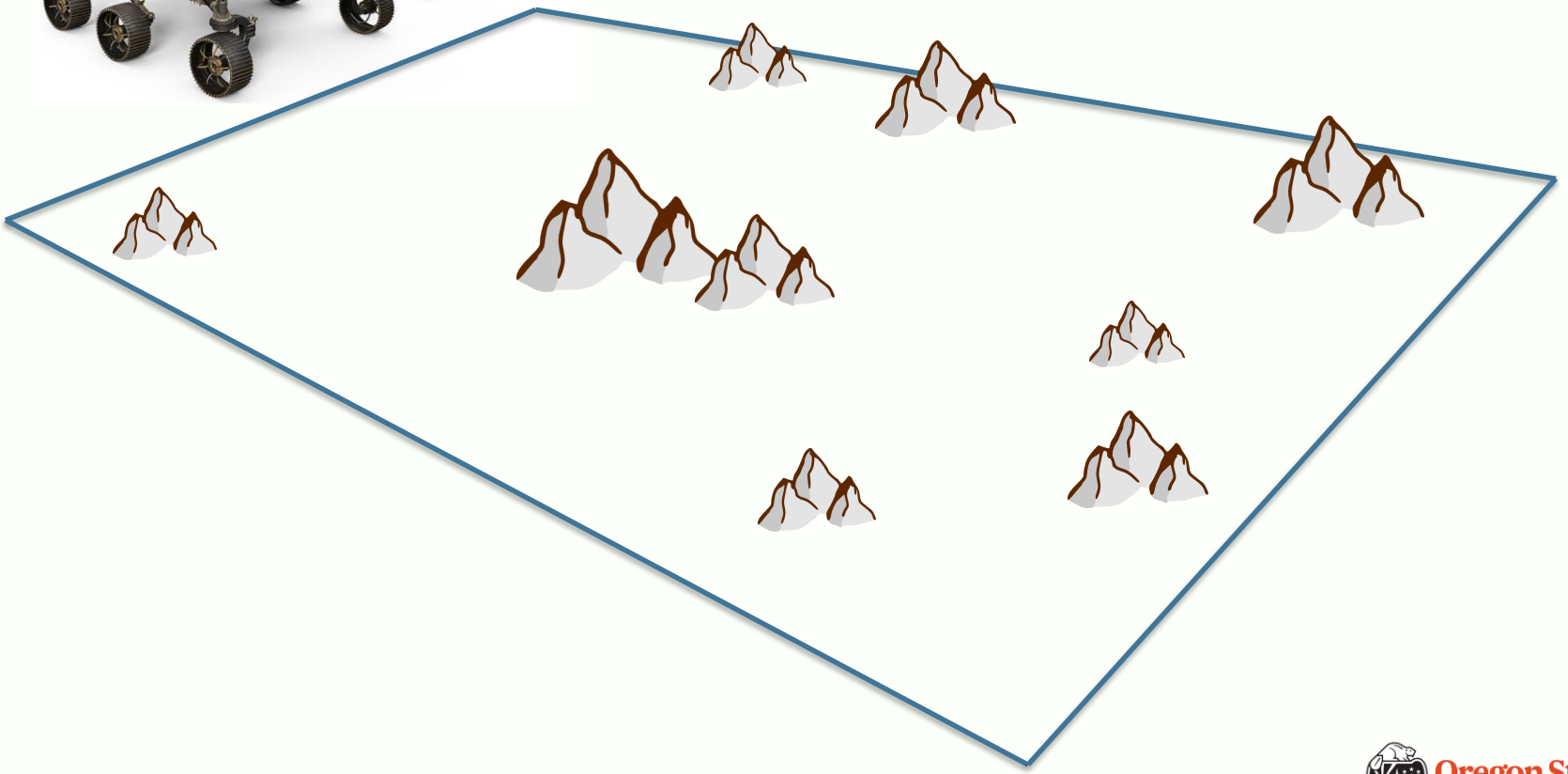
Have I seen this before ?

What matters when ?

Oregon State University

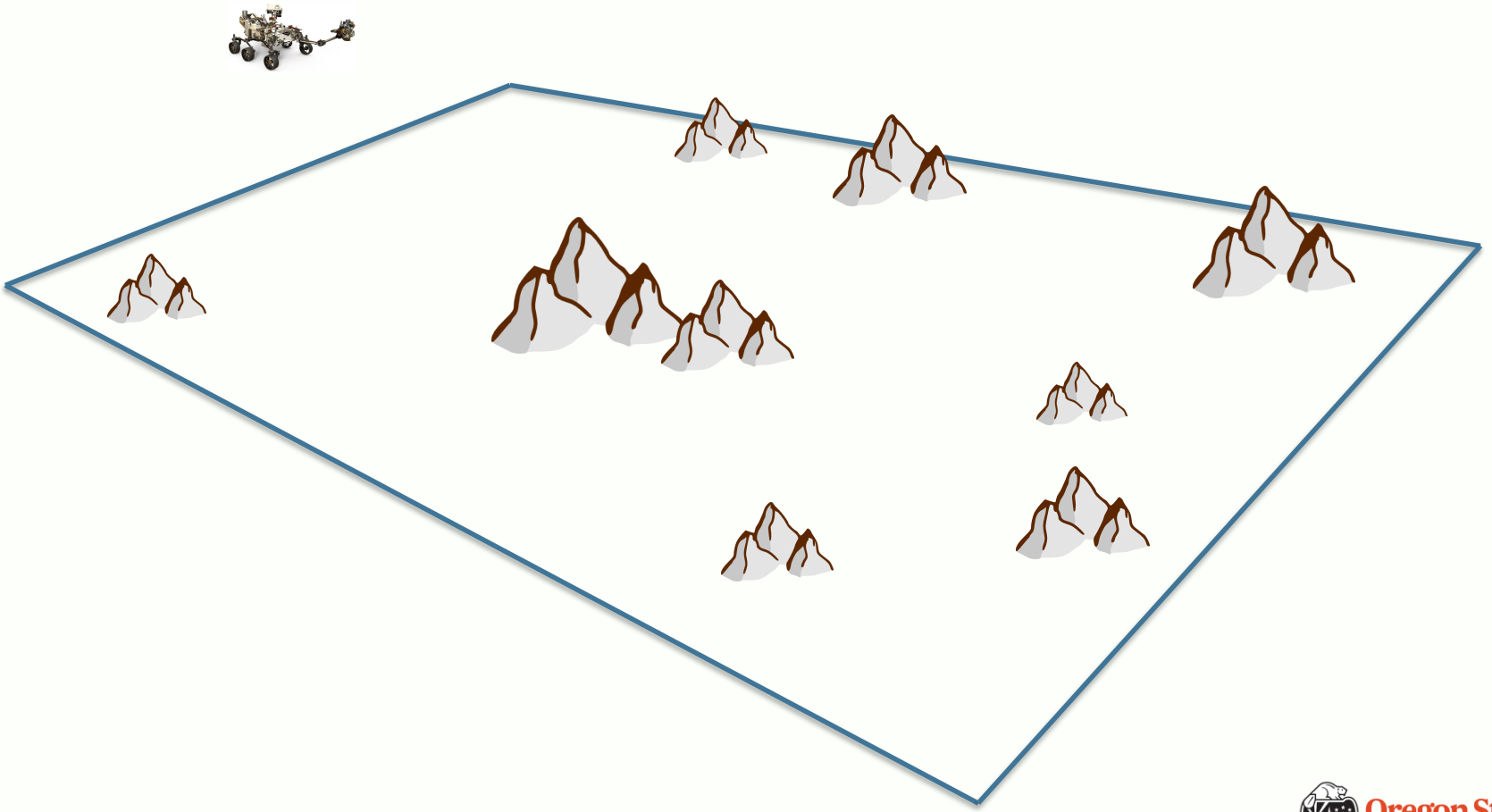
Domain: Robots in search of …





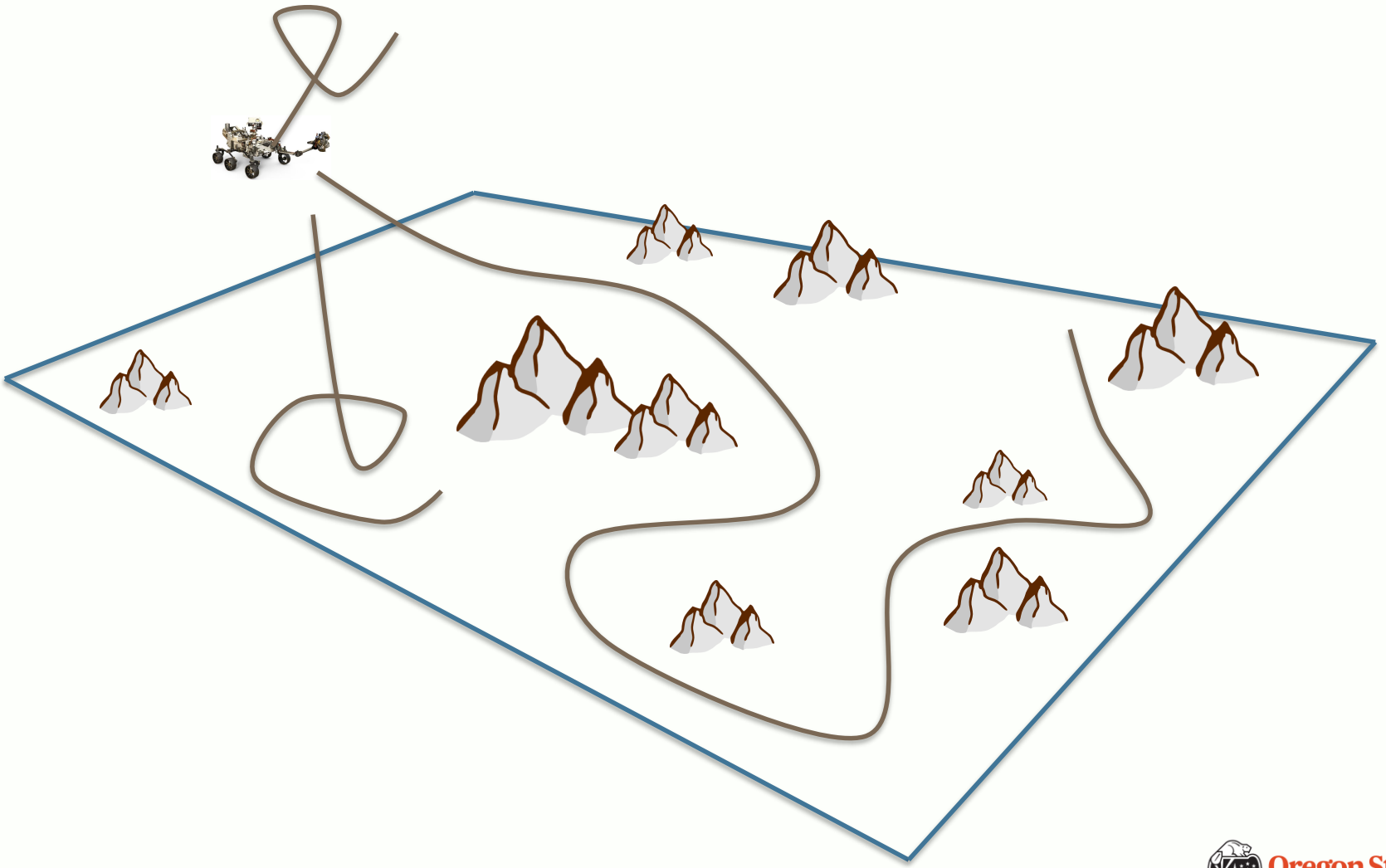
Oregon State University

Domain: Robots in search of …



Oregon State University

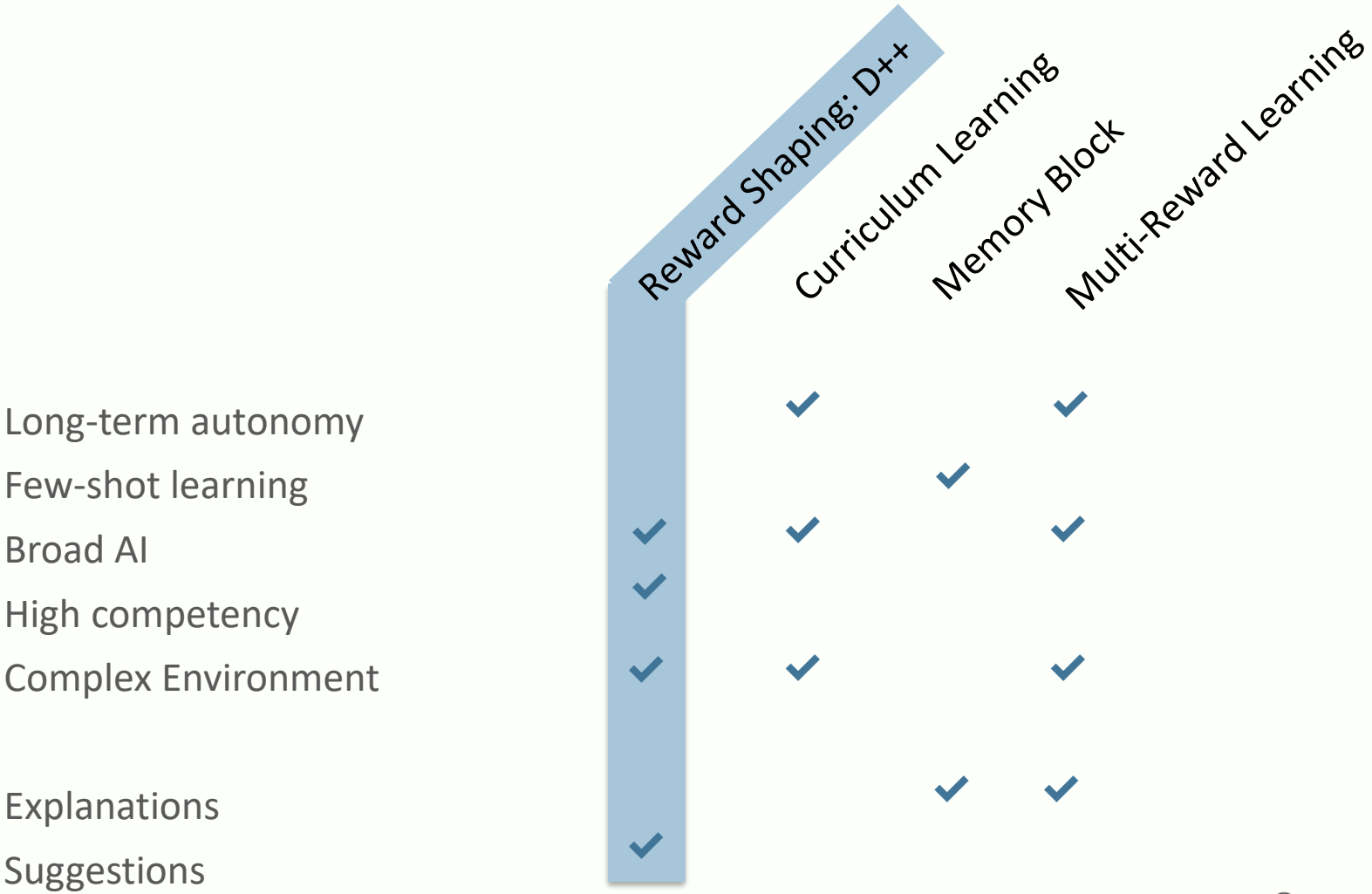
Domain: Robots in search of …



Oregon State University

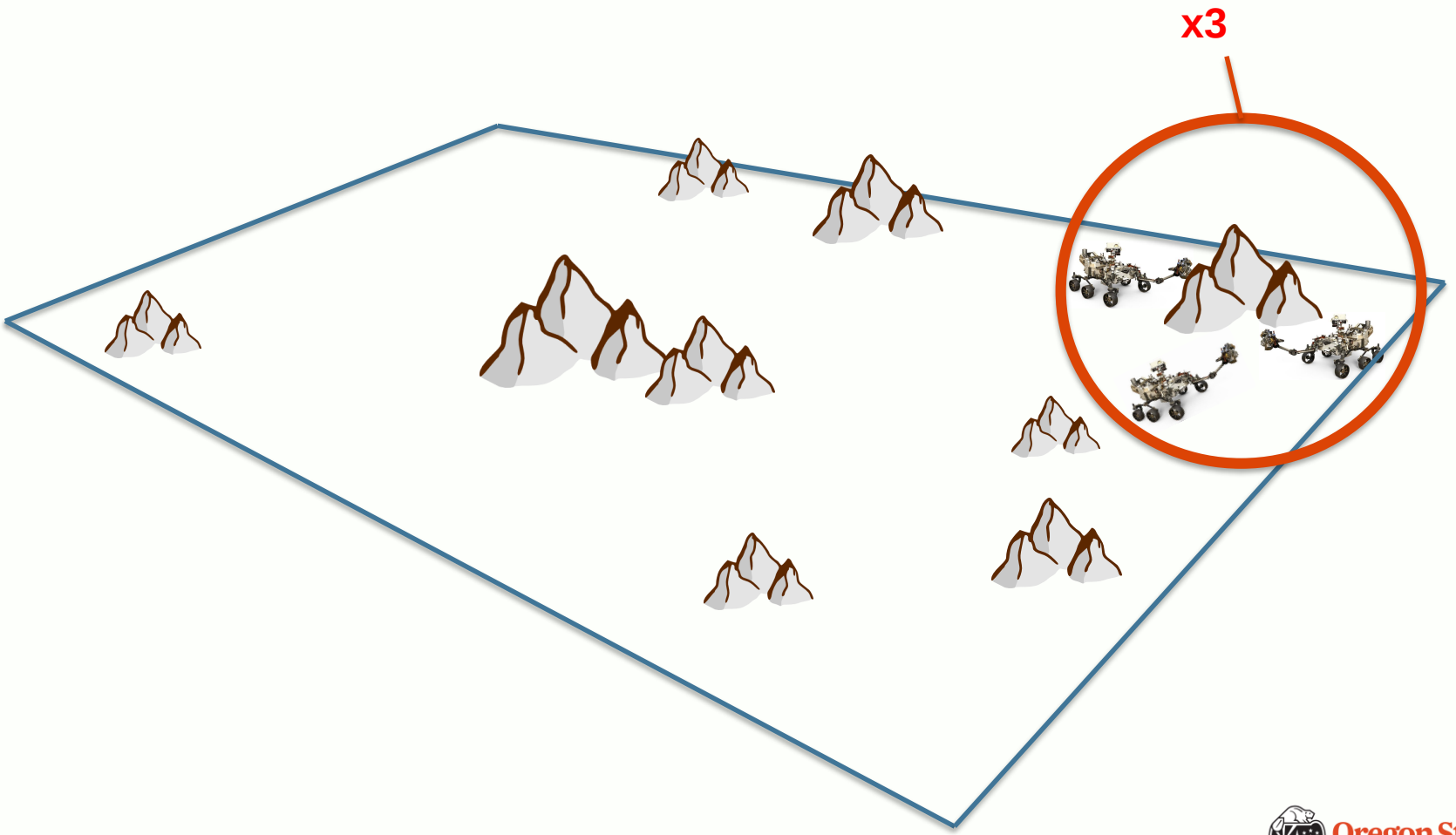
Motivating Example

What will it take to get where we're going ?



Oregon State University

Let’s look at a harder problem



Oregon State University

What happens when you need such close coordination ?

The probability of

SUFFICIENT agents, picking the **RIGHT ACTION**, at the **RIGHT TIME**

is **VANISHINGLY LOW**

Dirty secret underlying all learning:

You have to stumble upon the right action by accident

The right question

How can we devise agent-specific evaluation functions to reward the *stepping stone actions* towards success?

Difference Evaluation Function (Agogino and Tumer, 2004)

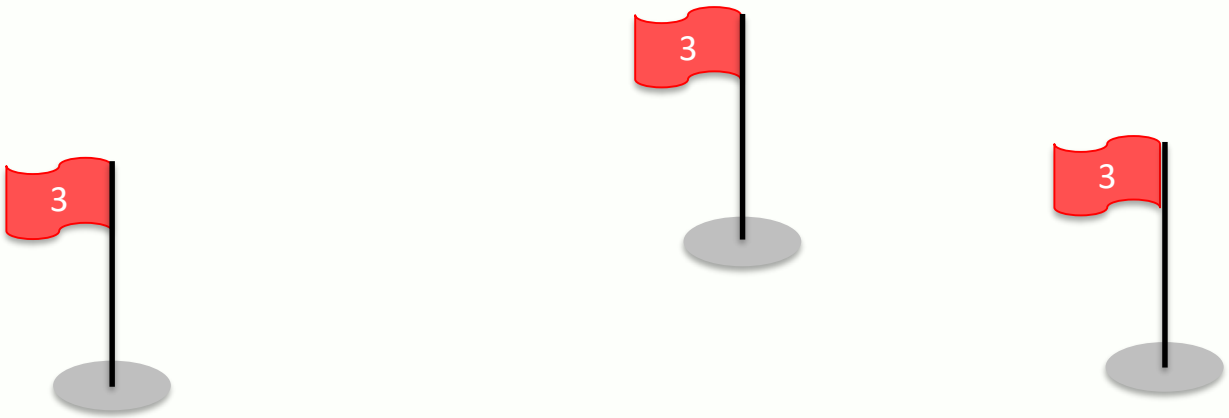
- Individual agents' contribution to the global team performance
- Removes an agent replaces a "*counterfactual*" agent

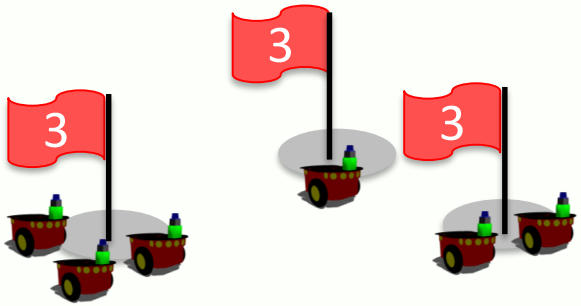
$$D_i(\vec{z}) = G(\vec{z}) - G(\vec{z}_{-i} \cup c_i)$$

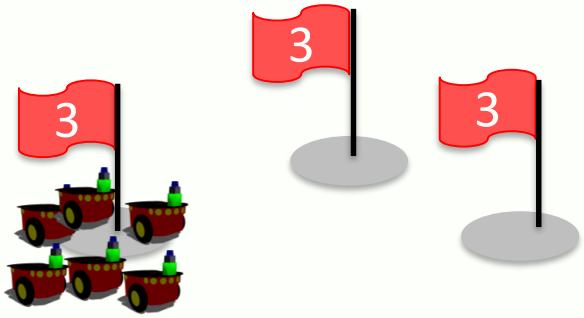
Global system performance
"The world with me"

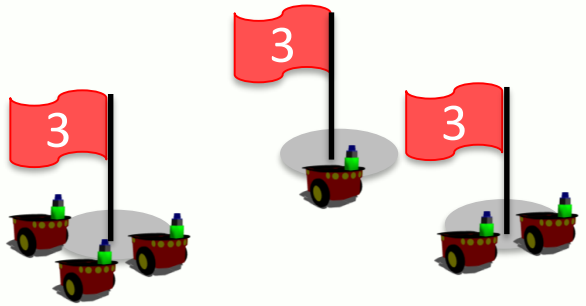
Global system performance
excluding the effects of agent *i*
"The world without me"

Oregon State University

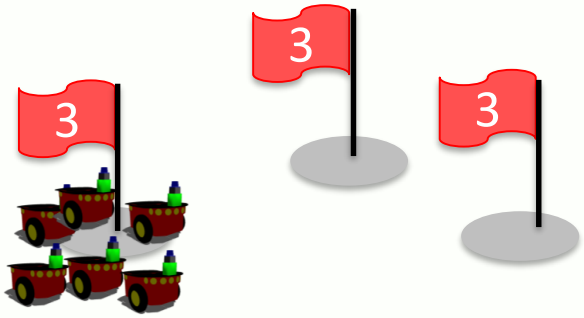




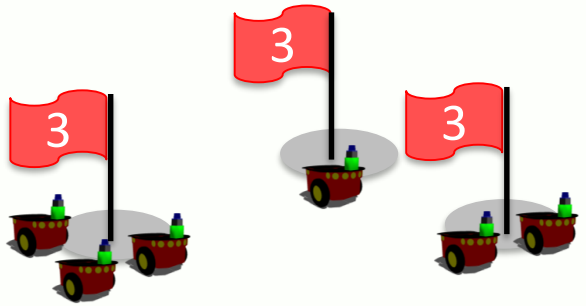




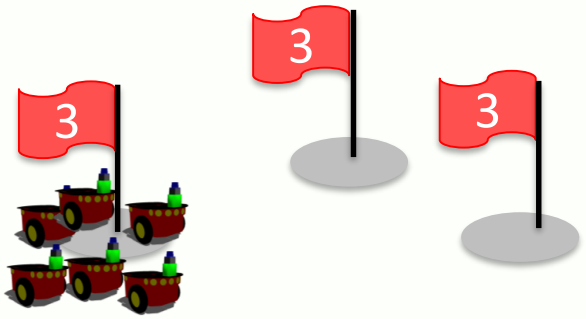
G	1	1	1



G	1	1	1



G	1	1	1
D	1	0	0



G	1	1	1
D	0	0	0

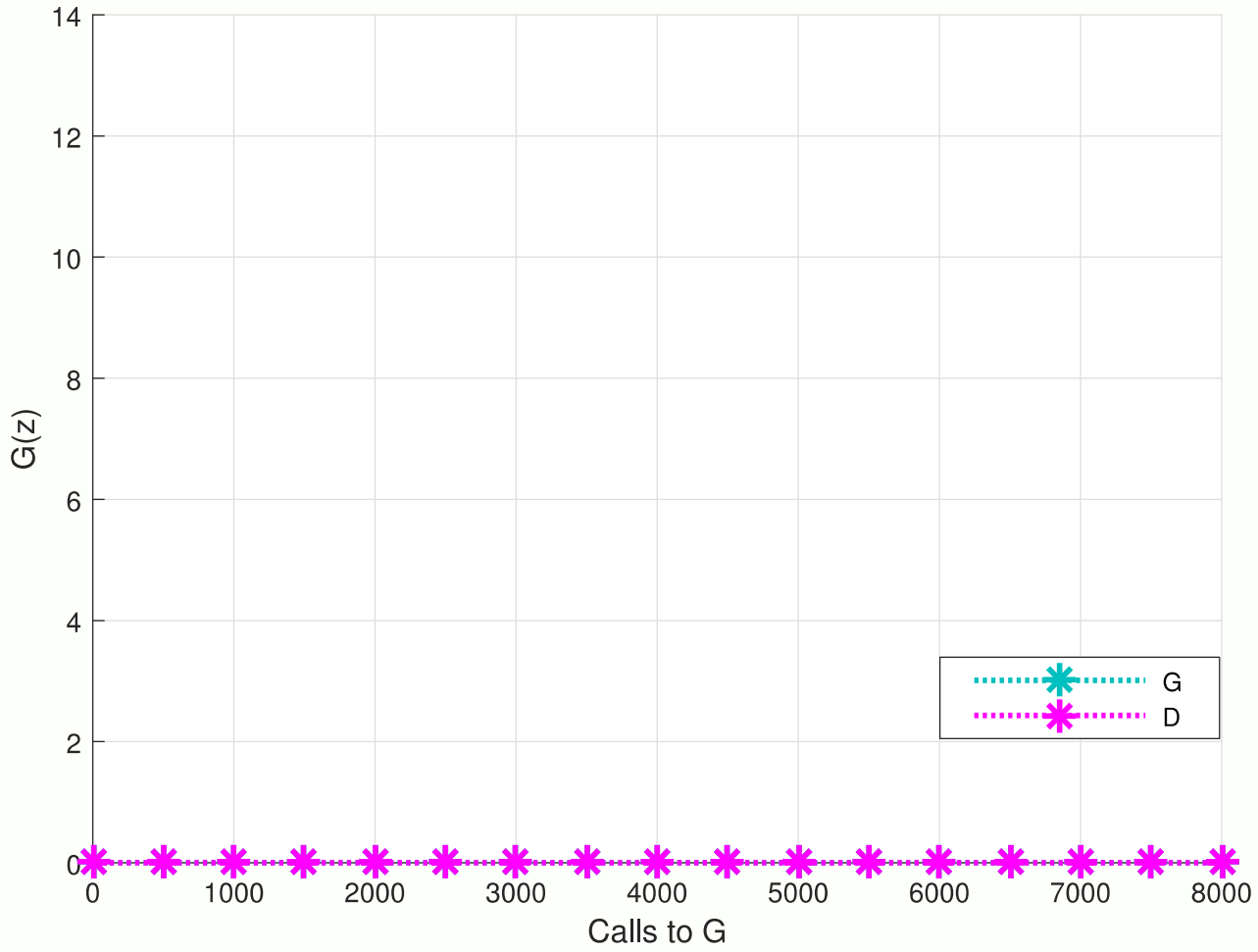
Problem

- Difference evaluation rewards good behavior
- BUT it does NOT differentiate between

  bad behavior and almost good behavior

Oregon State University

Rover Domain: 6 Observations Required



Oregon State University

D++: An Extension to Difference Reward (D)

– Think of "counterfactual agents" rather than counterfactual actions

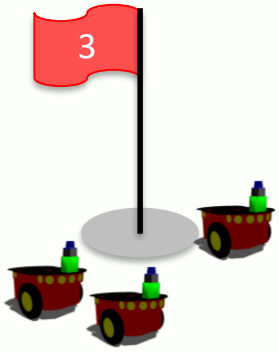
$$D^n_{++}(z) = \frac{G(\vec{z}_+(\cup_{i=1,\ldots,n})i) - G(\vec{z})}{n+1}$$

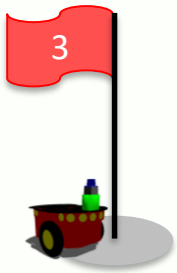
Global system performance
Where "multiple copies of me" are present

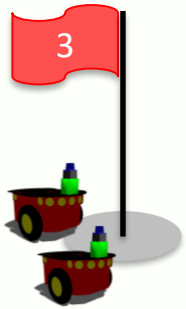
Global system performance

– Provides agents with stronger feedback signal
– Rewards the stepping stones that lead to achieving the system objective
– Actual Algorithm: Pick max of D and D++

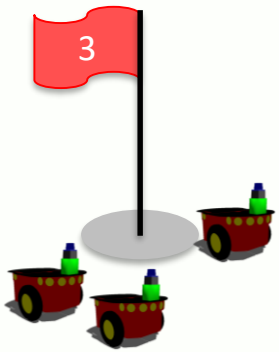
Oregon State University

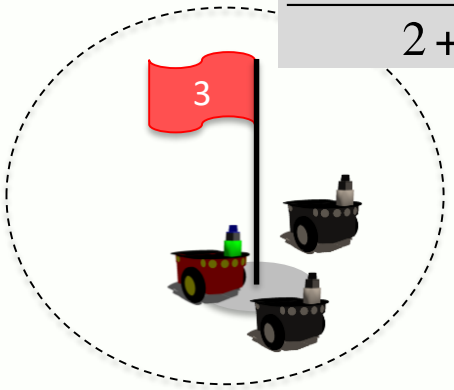


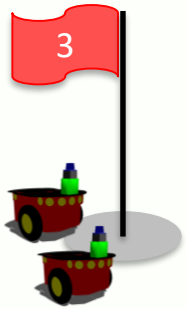


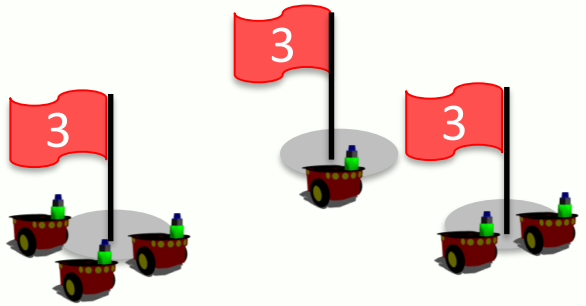


$$\frac{G(\vec{z} \cup_{2i}) - G(\vec{z})}{2+1}$$

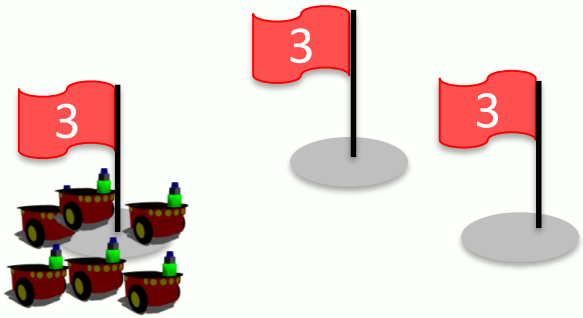




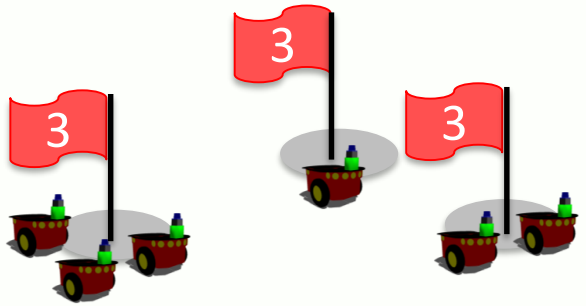




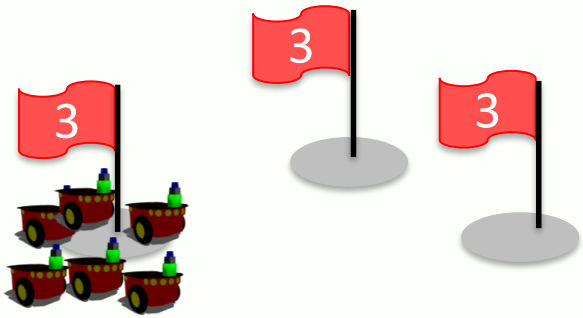
G	1	1	1
D	1	0	0
D++	1	0.33	0.5



G	1	1	1
D	0	0	0
D++	0	0	0

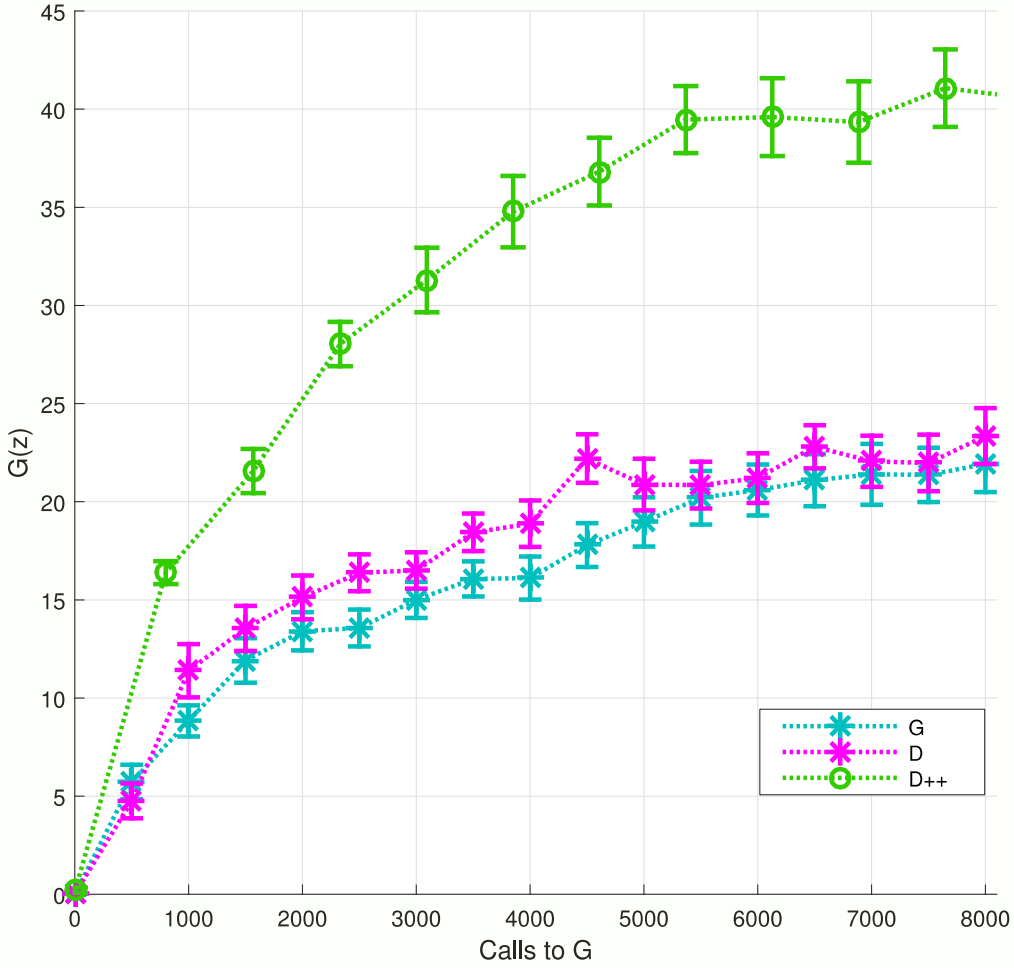


G	1	1	1
D	1	0	0
D++	1	0.33	0.5



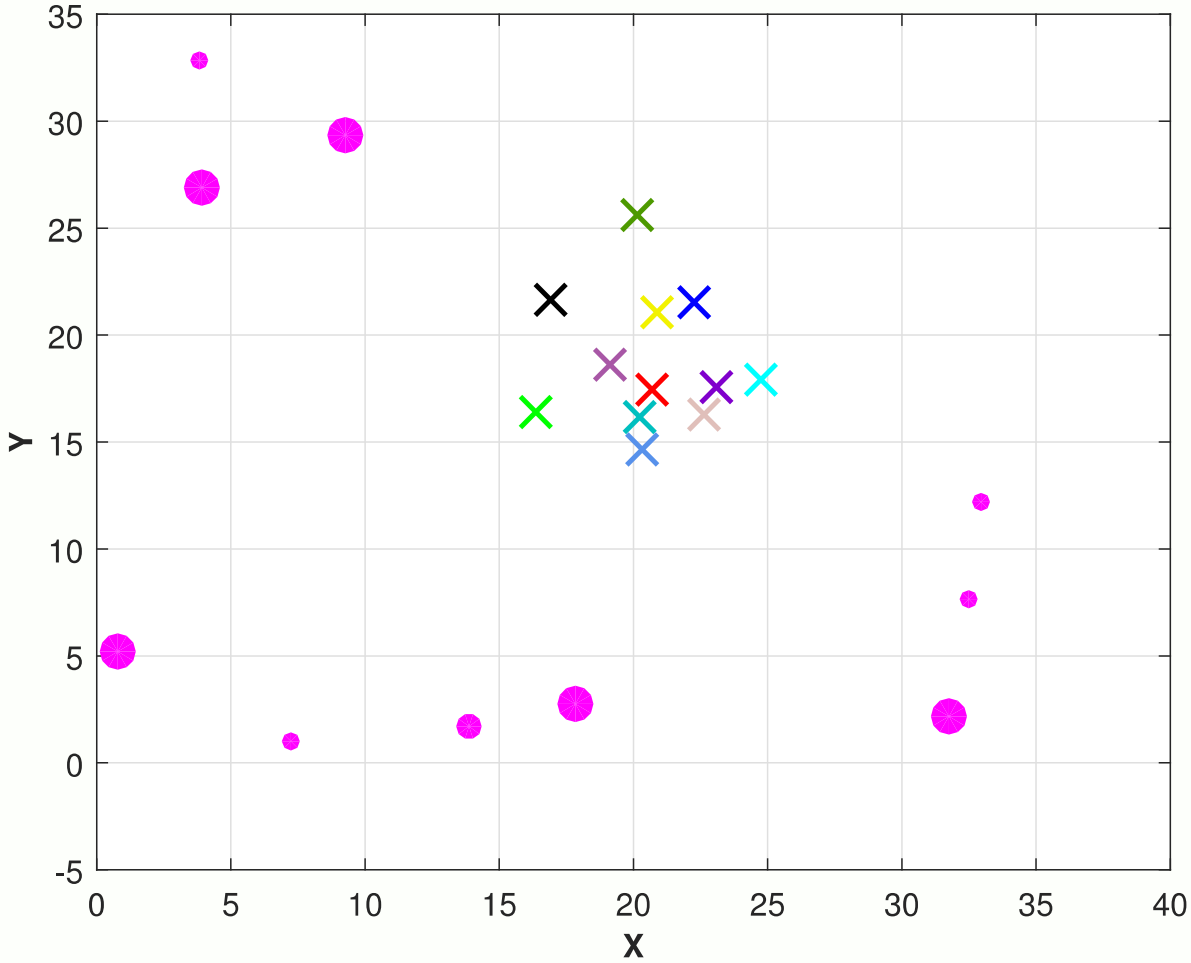
G	1	1	1
D	0	0	0
D++	0	0	0

Rover Domain: 3 Observations Required



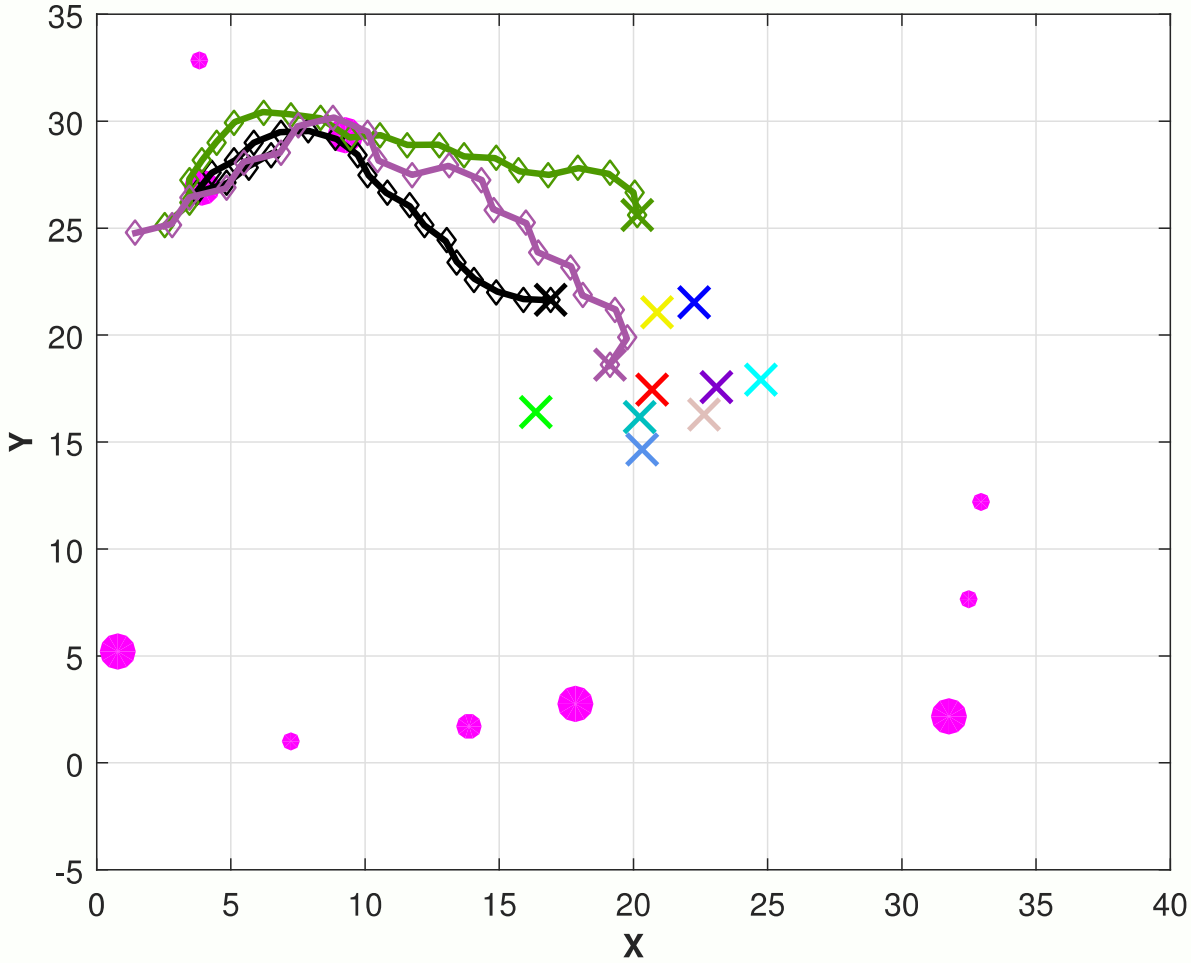
Oregon State University

Rover Domain: Learned Policies of D++ learners



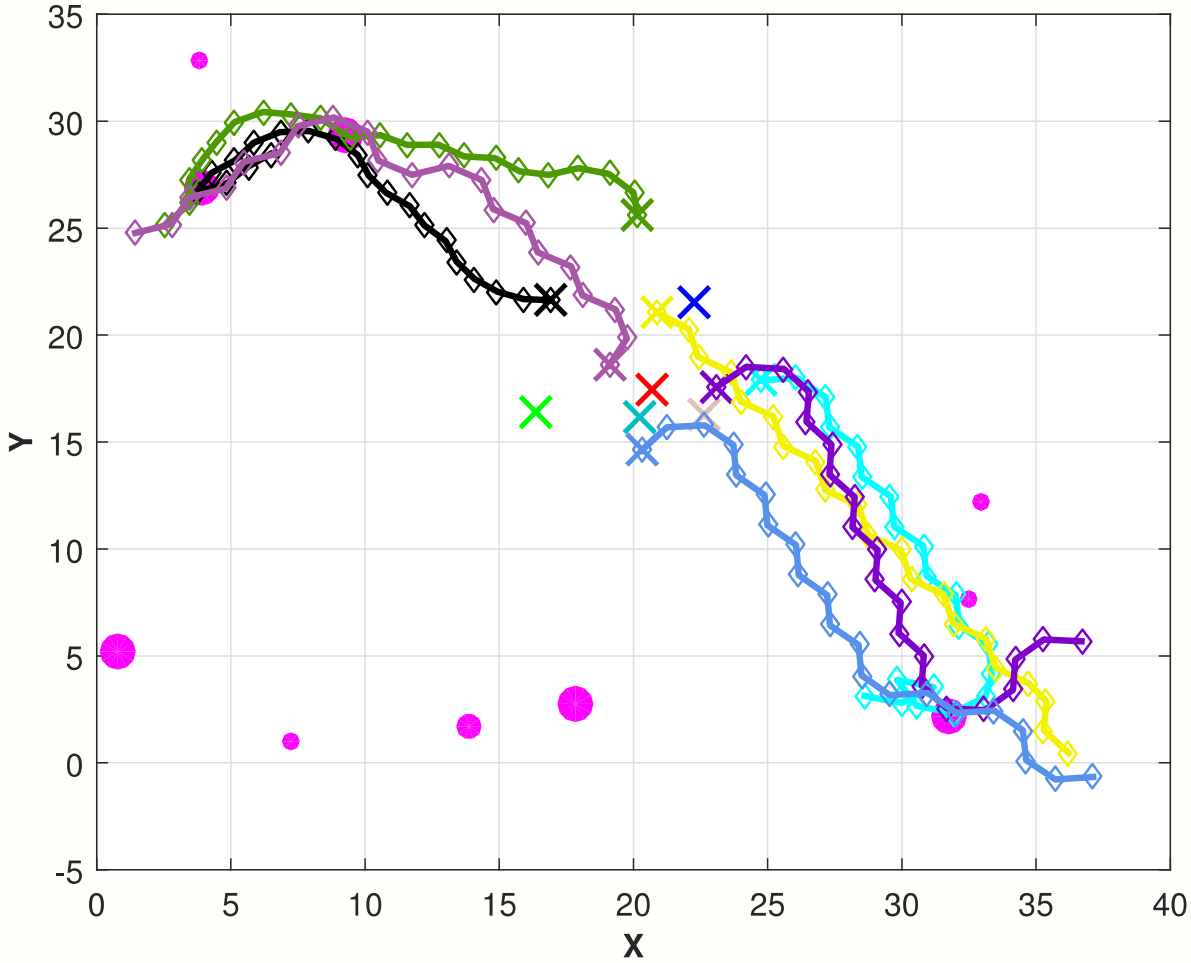
Oregon State University

Rover Domain: Learned Policies of D++ learners



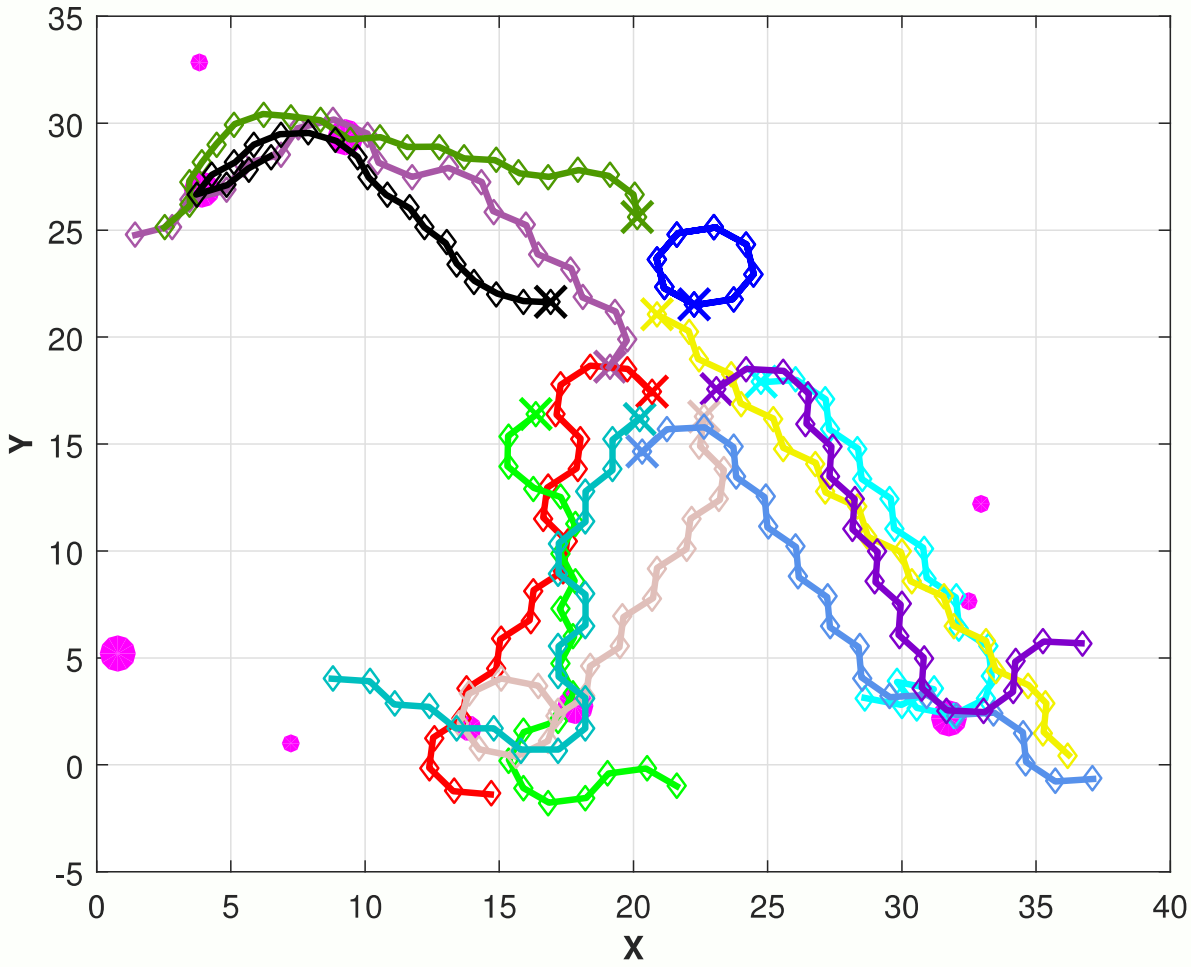
Oregon State University

Rover Domain: Learned Policies of D++ learners



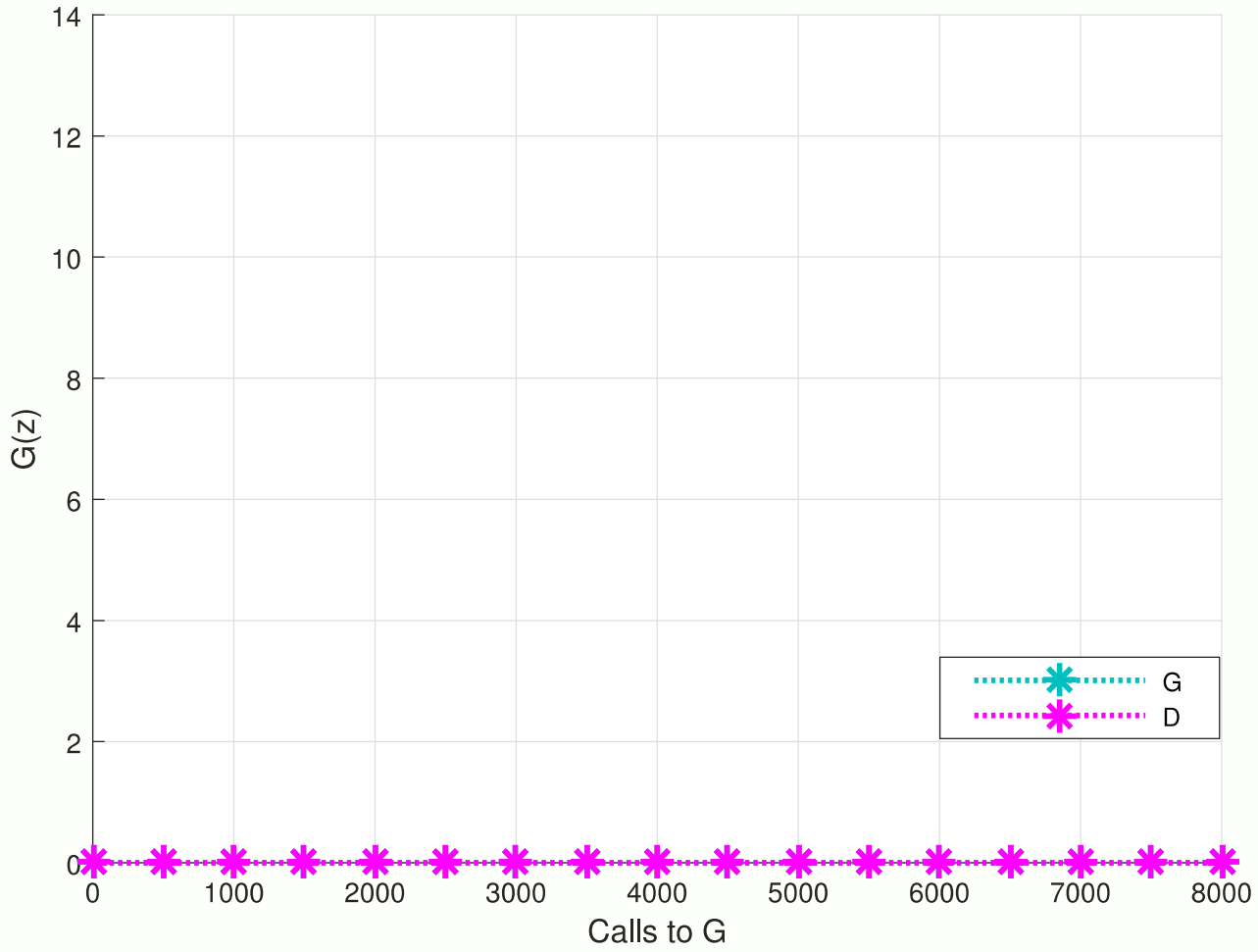
Oregon State University

Rover Domain: Learned Policies of D++ learners



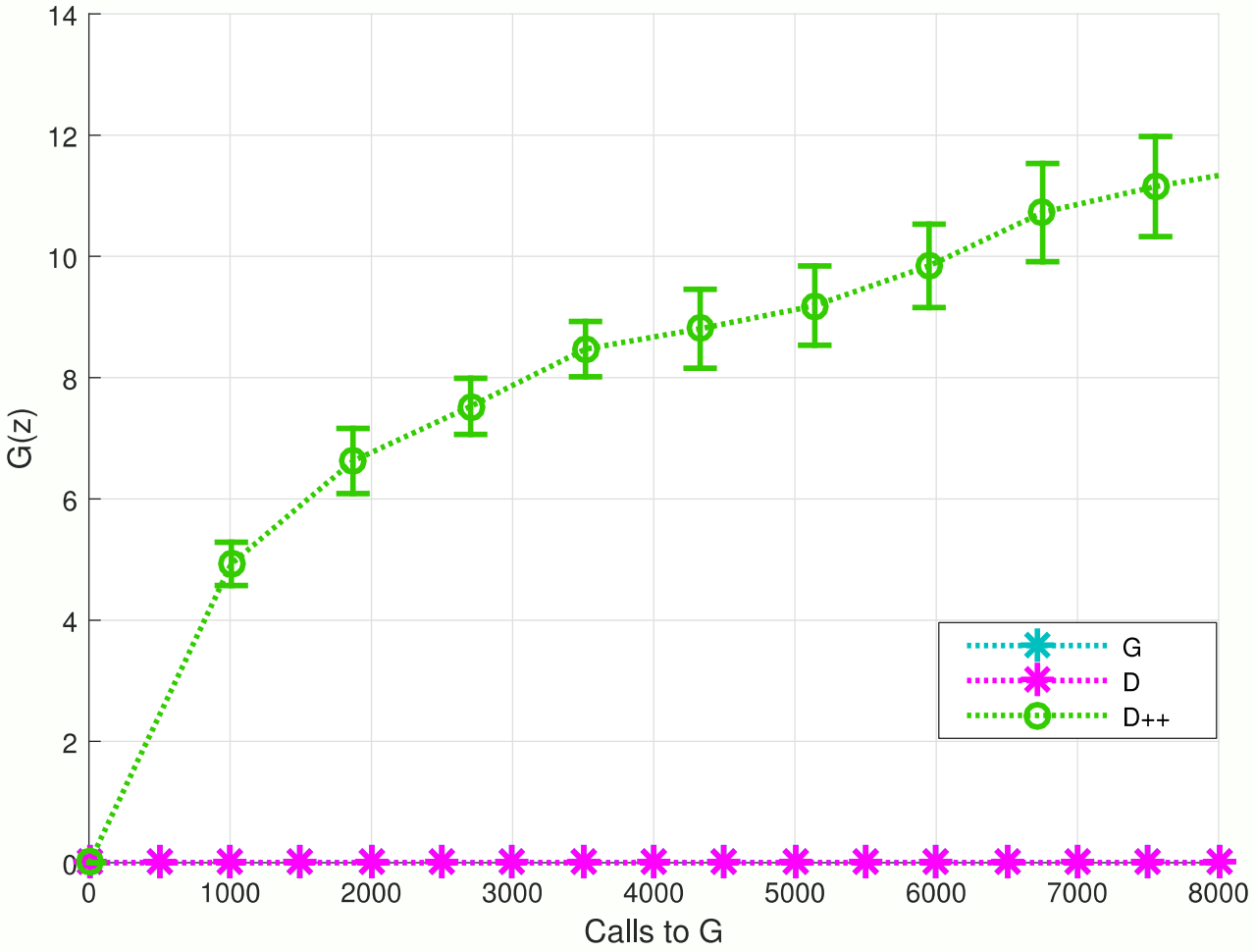
Oregon State University

Rover Domain: 6 Observations Required



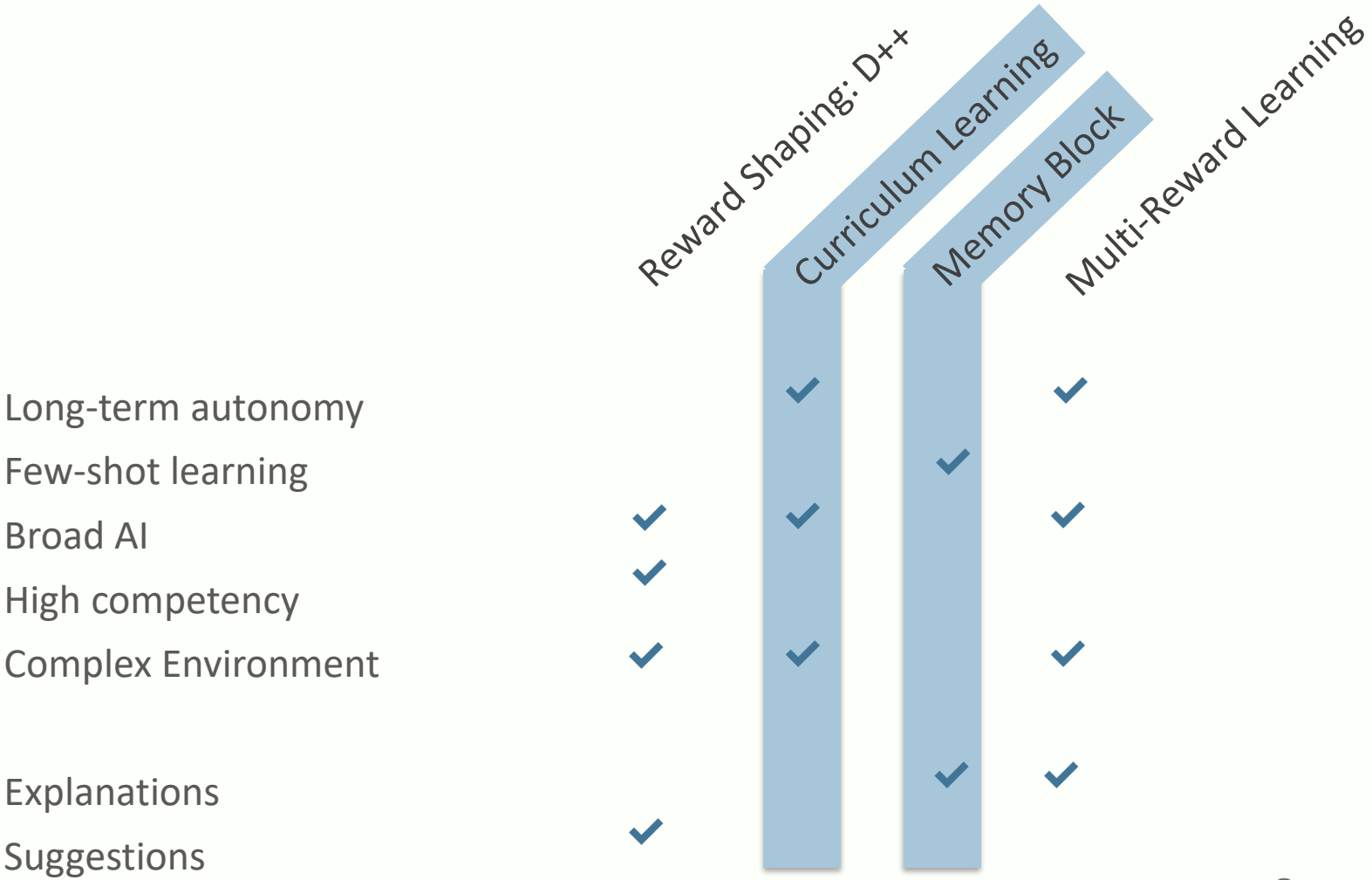
Oregon State University

Rover Domain: 6 Observations Required



Oregon State University

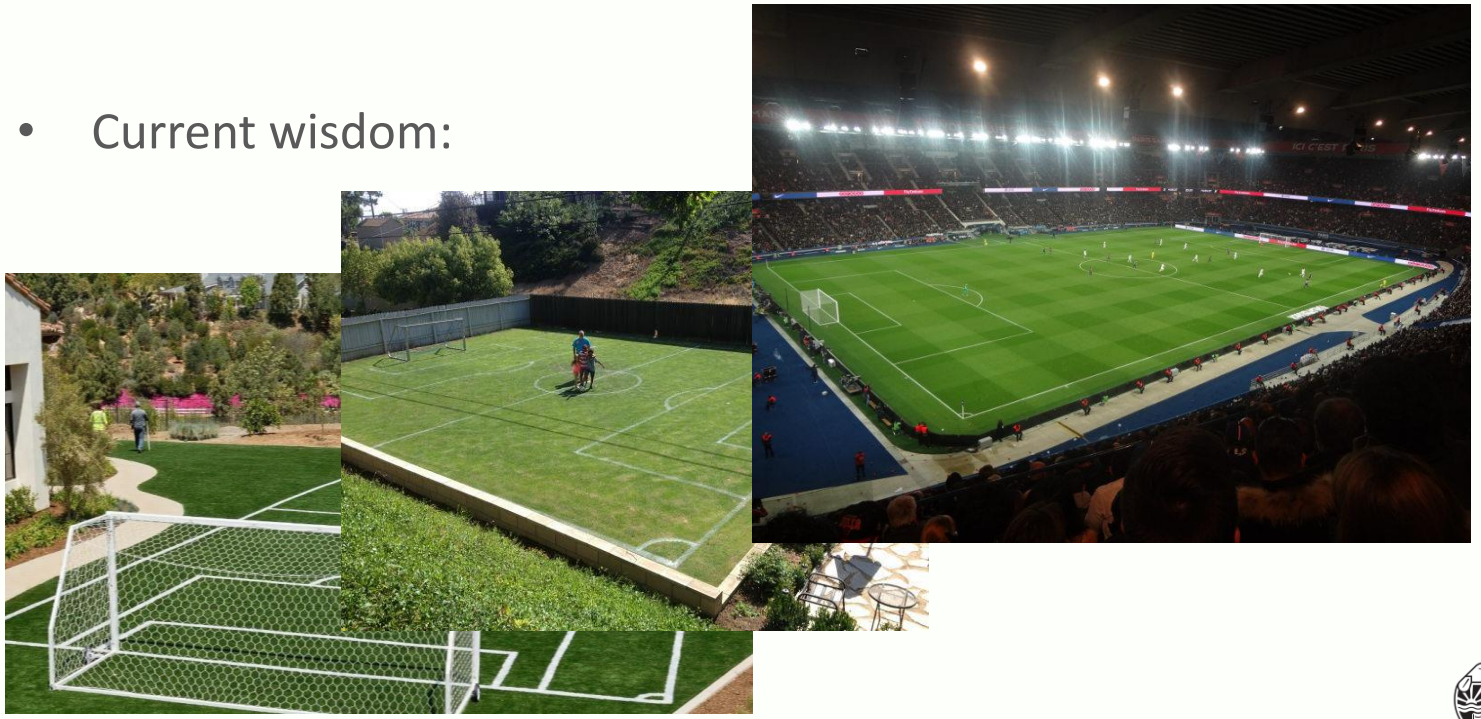
What will it take to get where we're going ?



Oregon State University

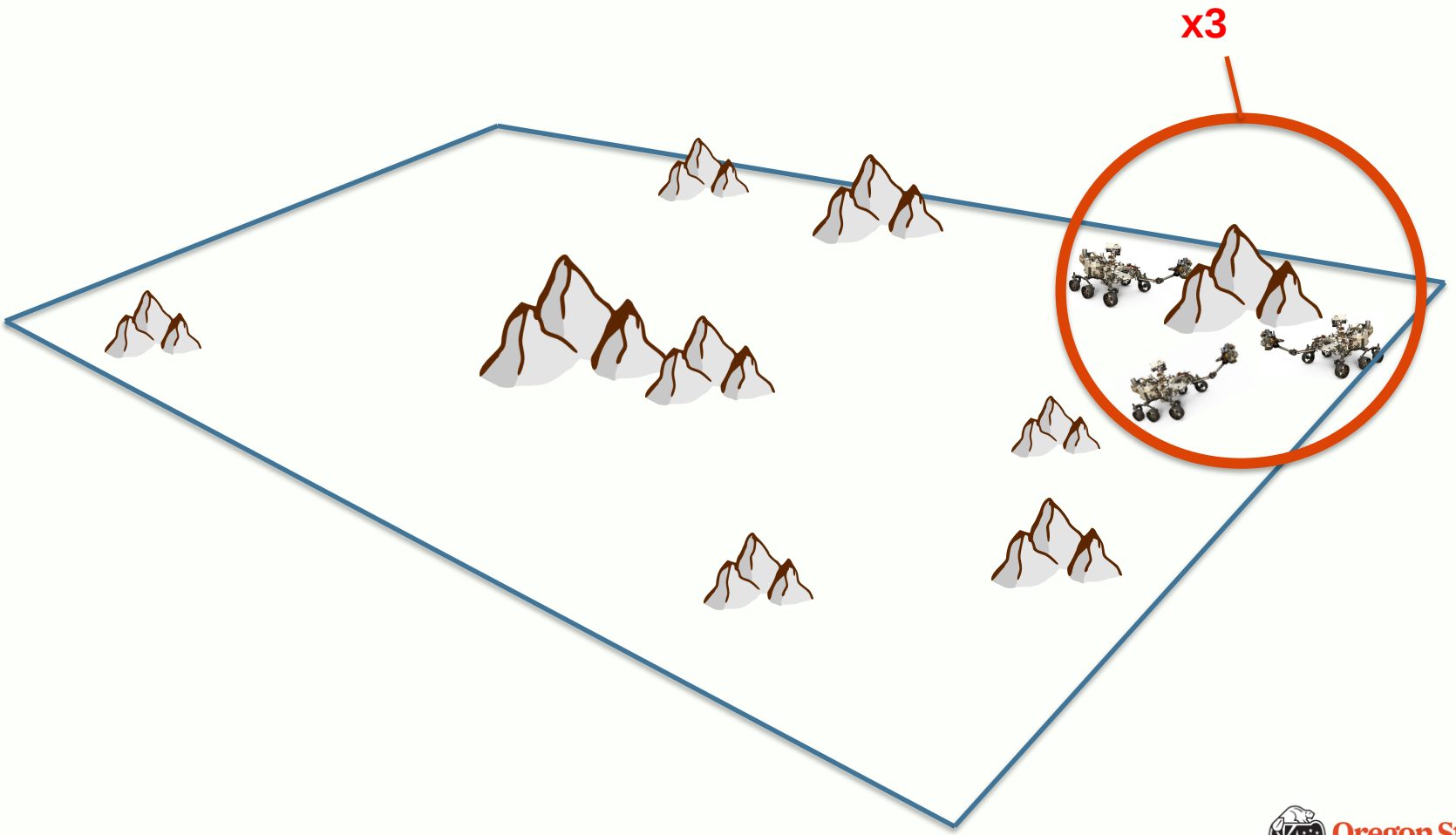
Curriculum Learning

- Learn to play soccer from winning game
- What game?
- Current wisdom:

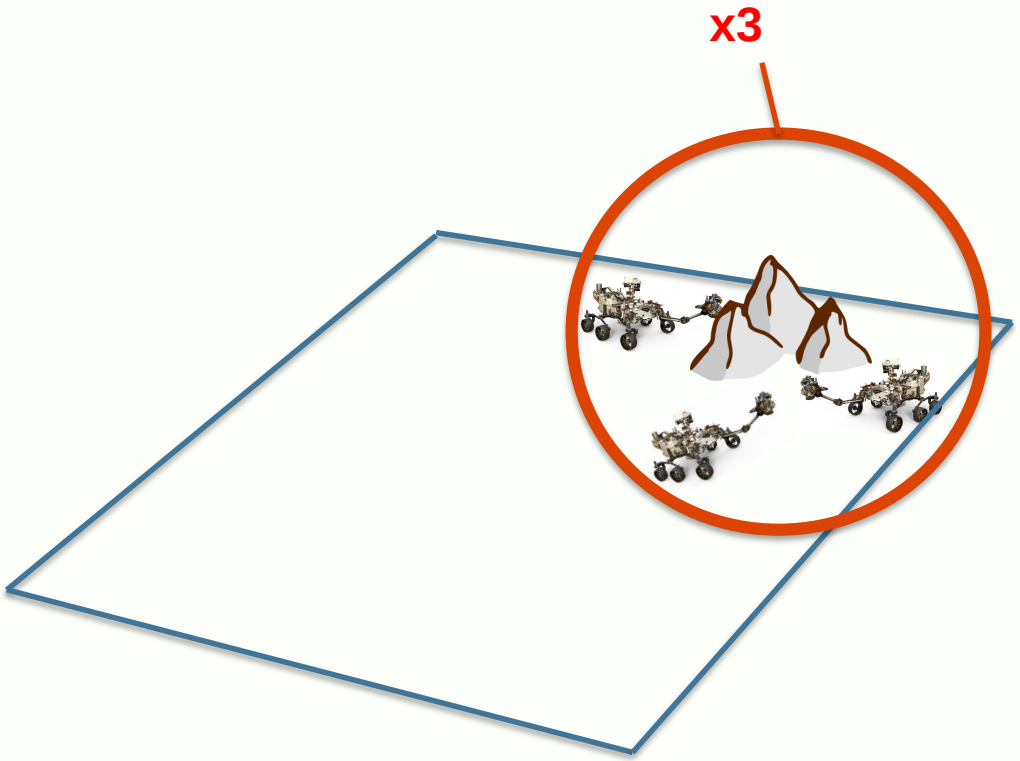


Oregon State University

Recall last problem

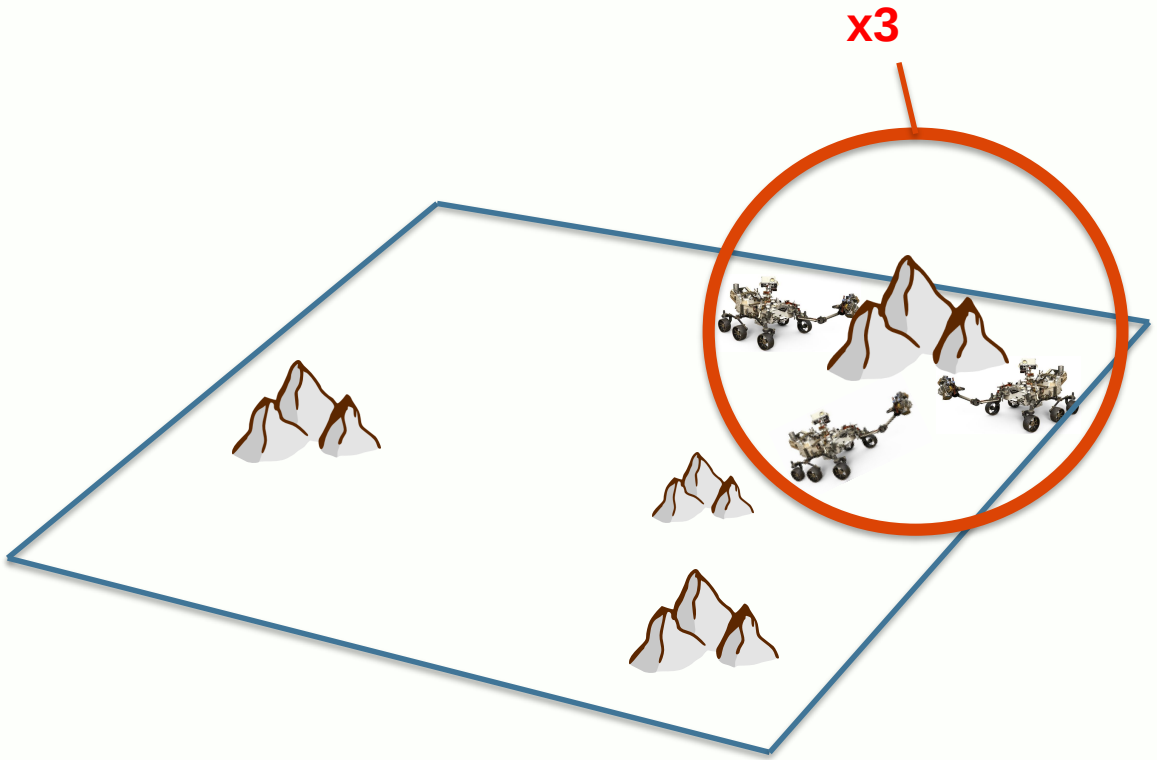


Simplify and learn



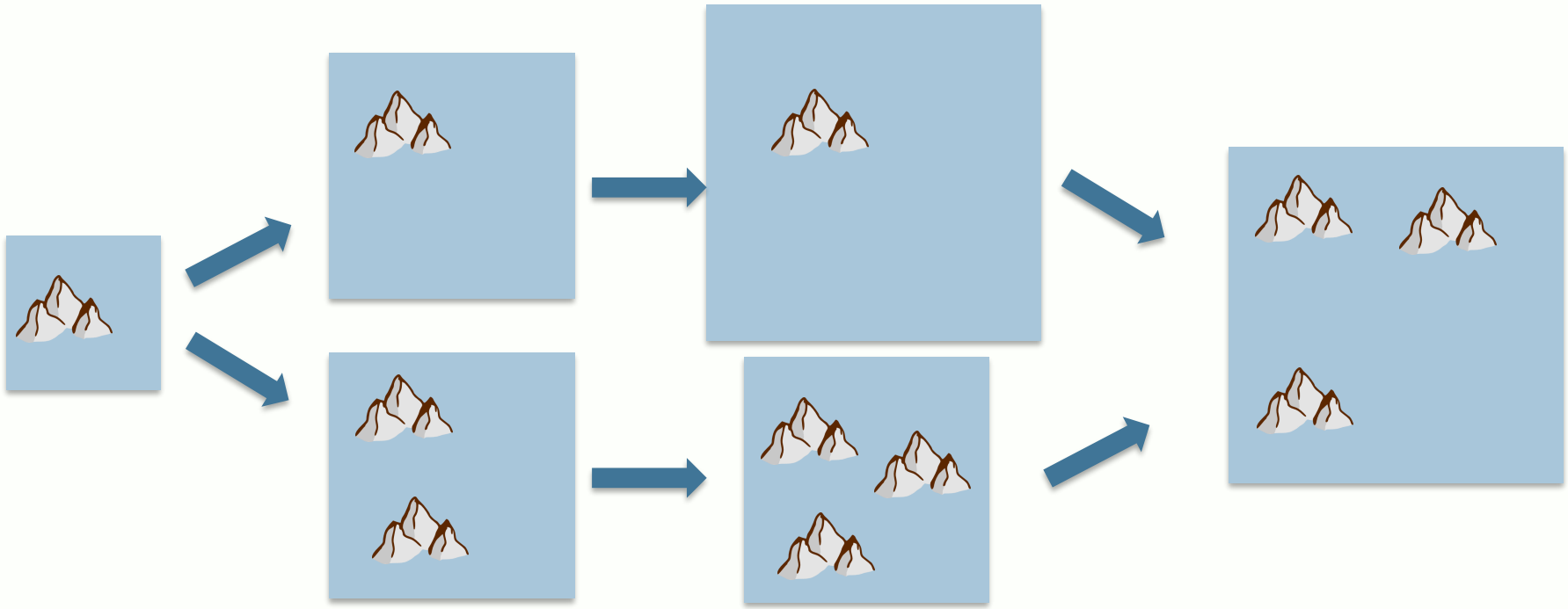
Oregon State University

Add complexity



Oregon State University

Design Curriculum



- Schedule of how space, task, action complexity increases is critical
- How to find the right curriculum?

Oregon State University

How to generate a curriculum ?

Ask Golden:

Oregon State University

Memory

Behave differently in different situations

Behave differently with different teammates

Oregon State University

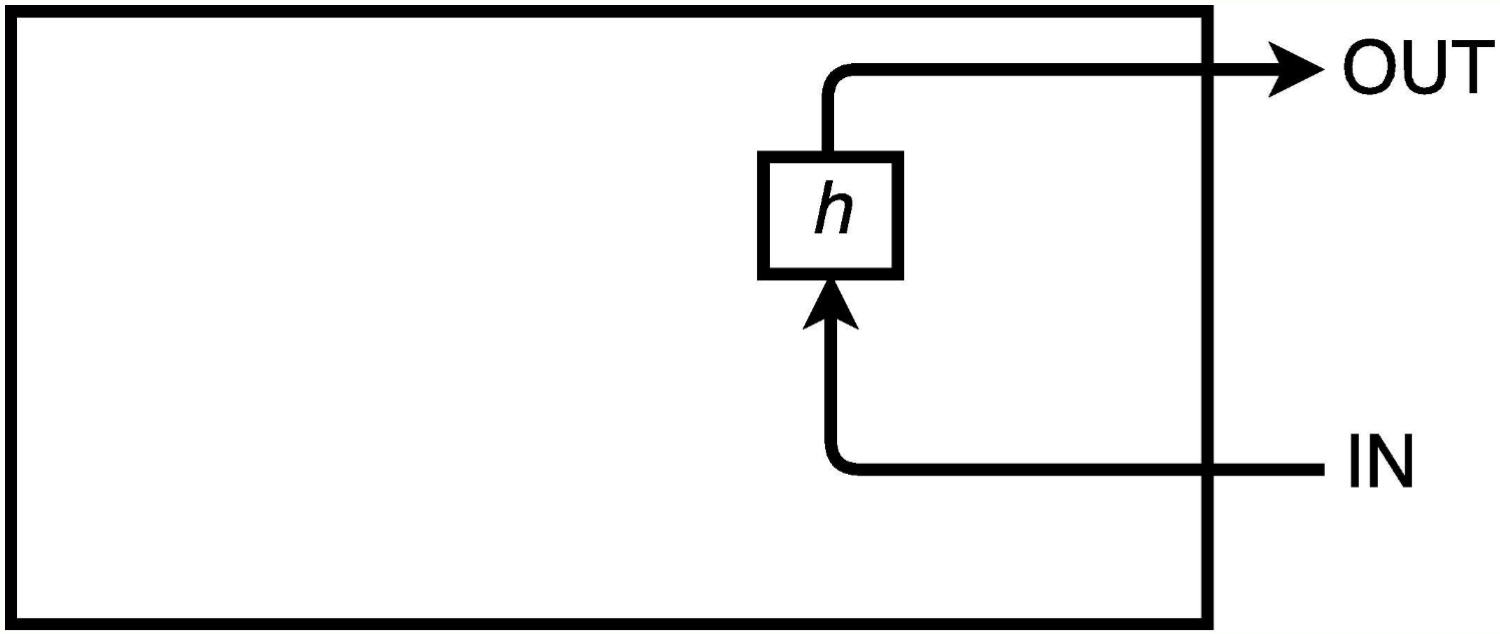
Gated Recurrent Unit with Memory Block (GRU-MB)

- Detach memory from action (arbitrary read/write)

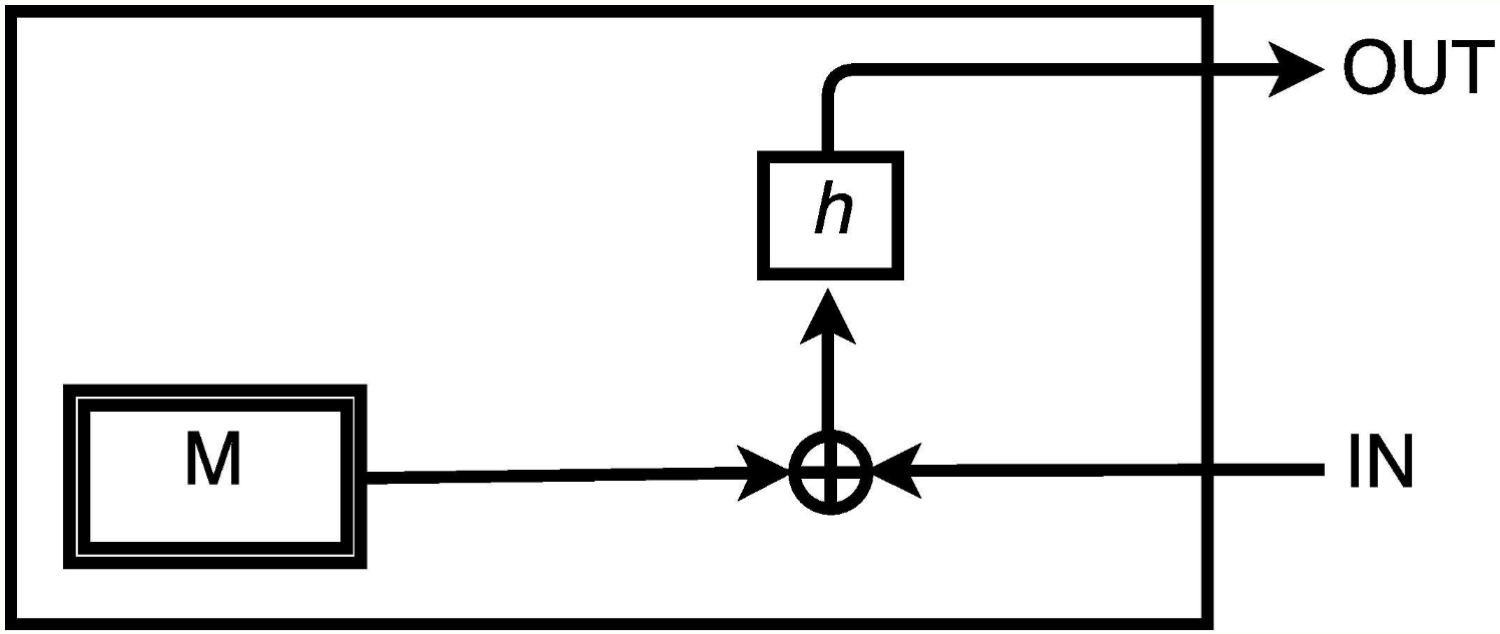
- similar to LSTM, but with key advantage

    - Ease of training
    - **Arbitrary access to memory**
    - **Decoupling of memory from action**

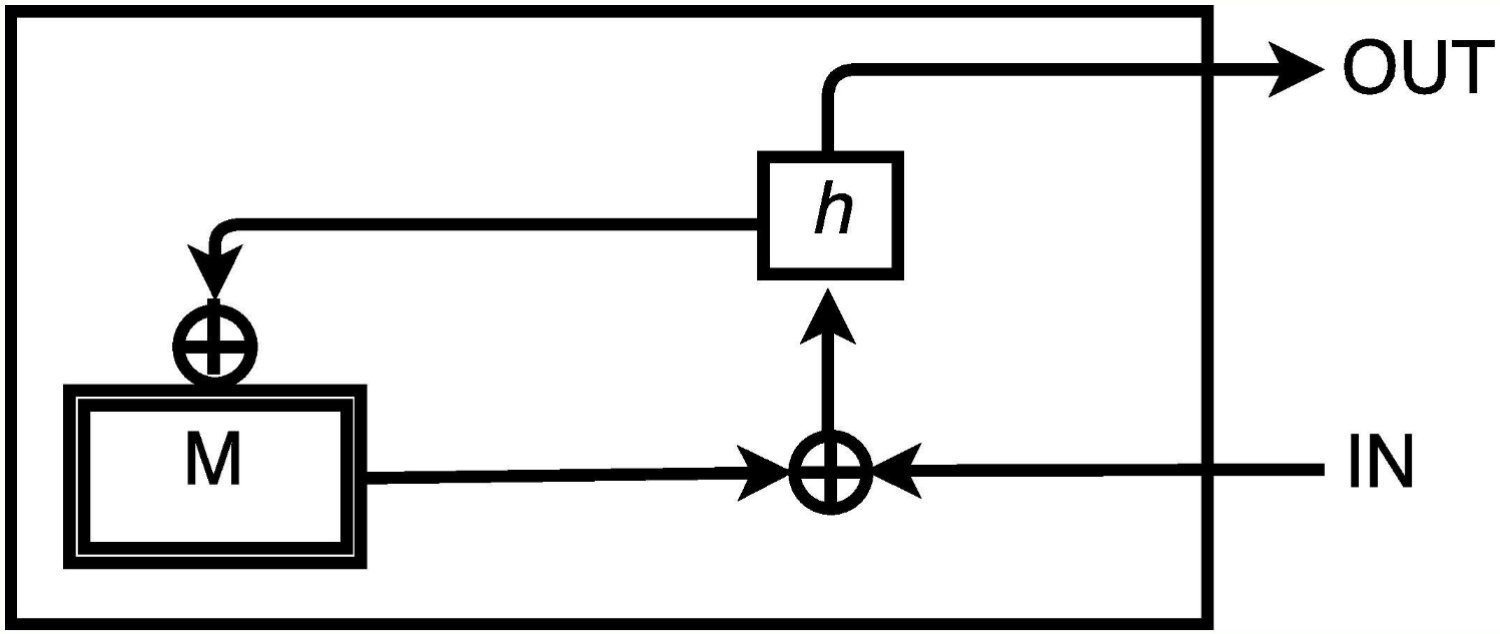
Feedforward Neural Network



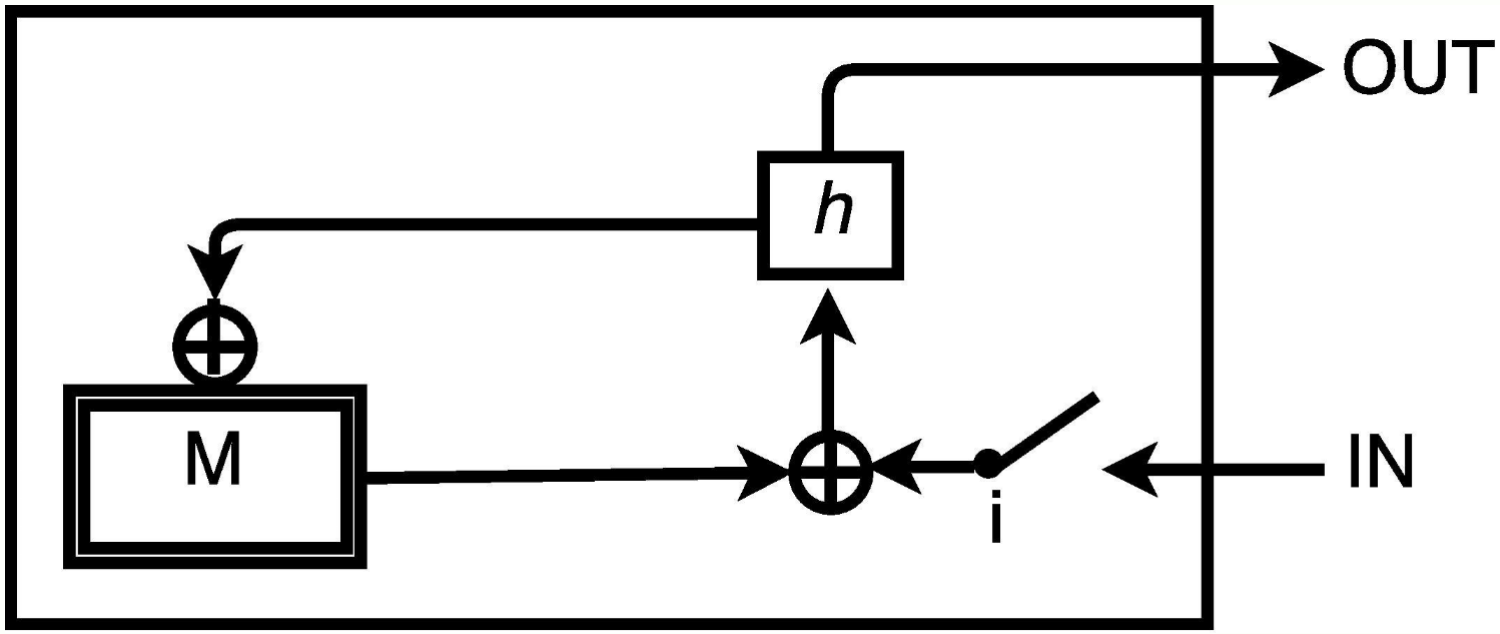
Read from external memory



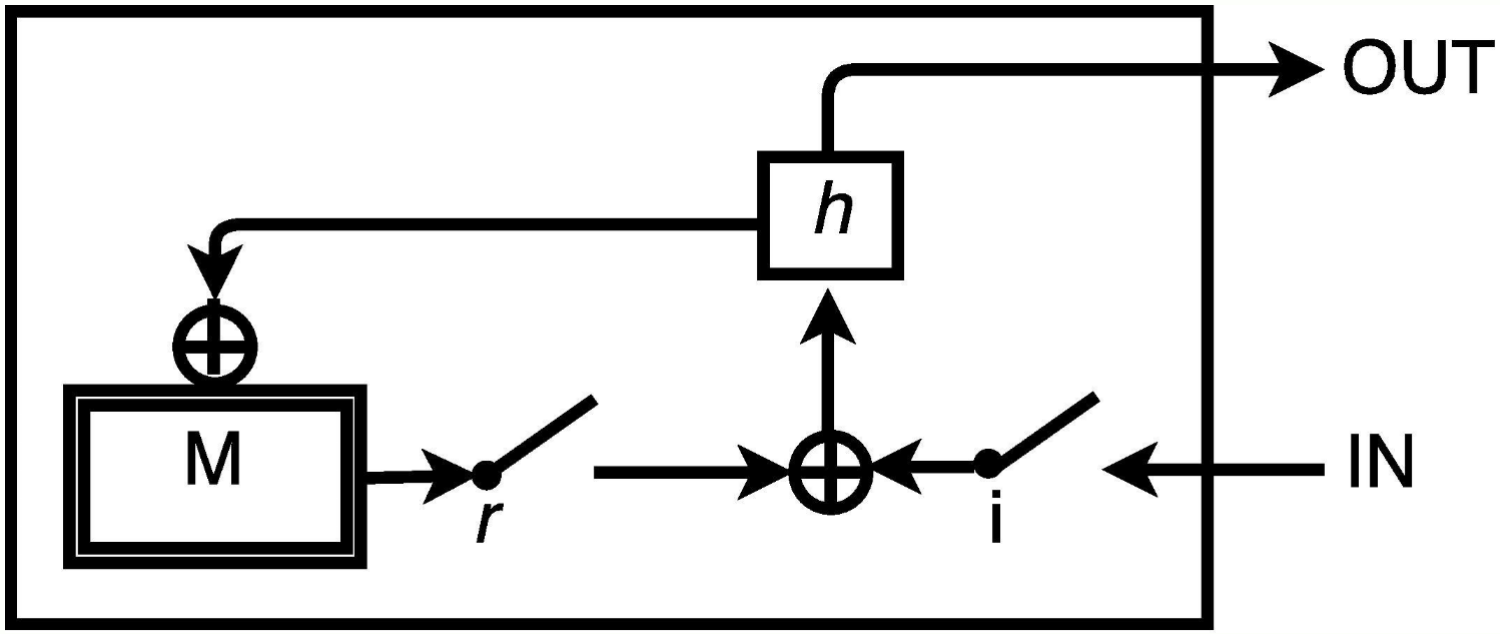
Write to memory



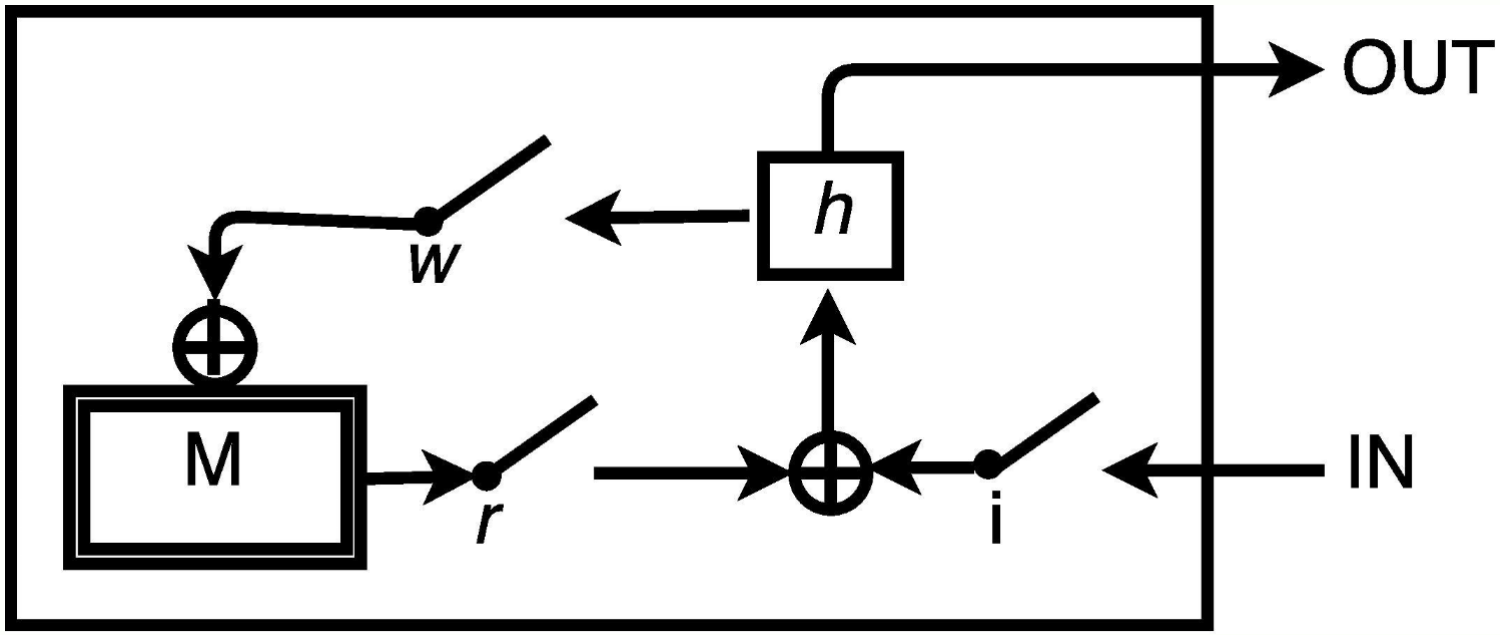
Gate Input



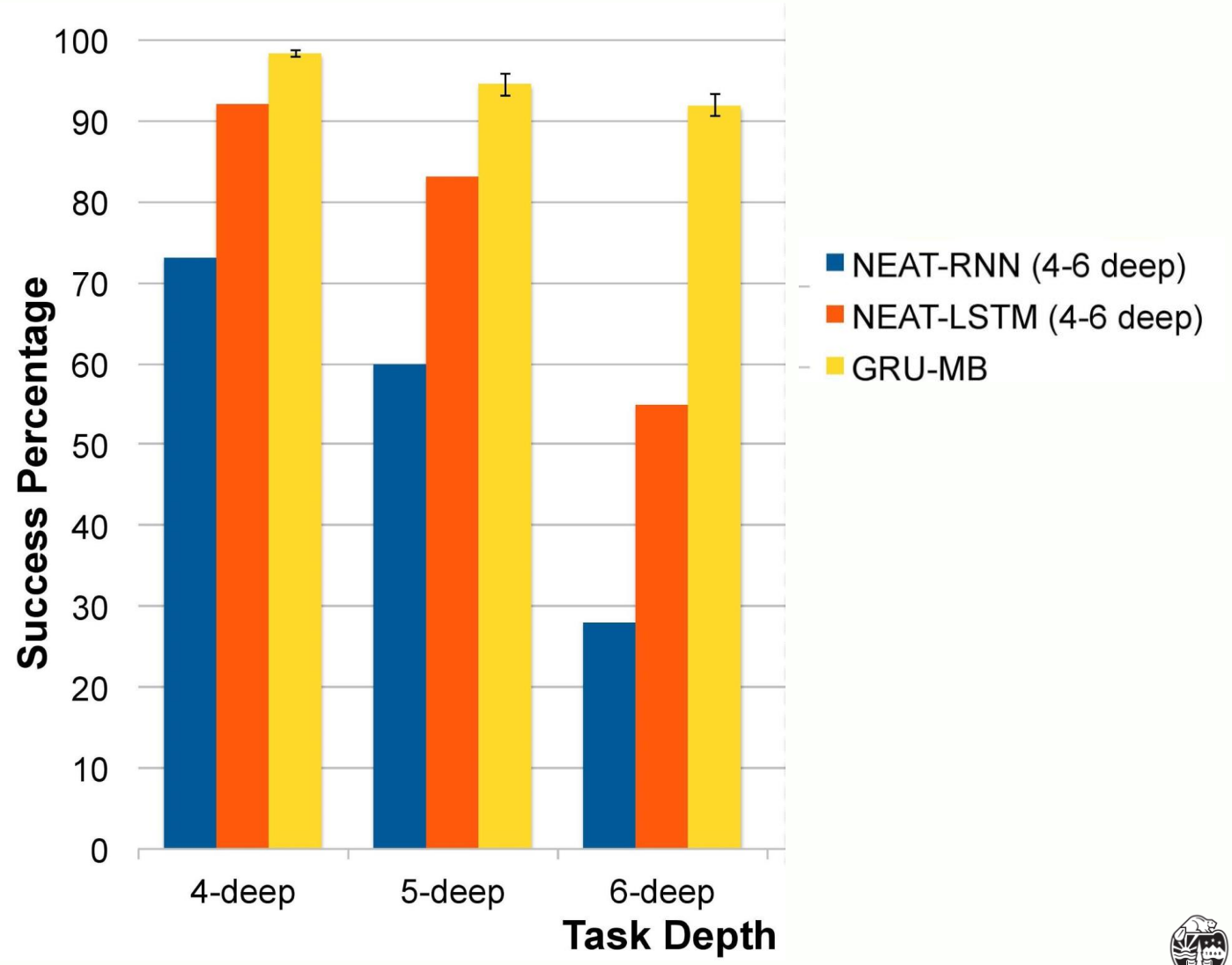
Gate what's read from memory



Gate what's written to memory

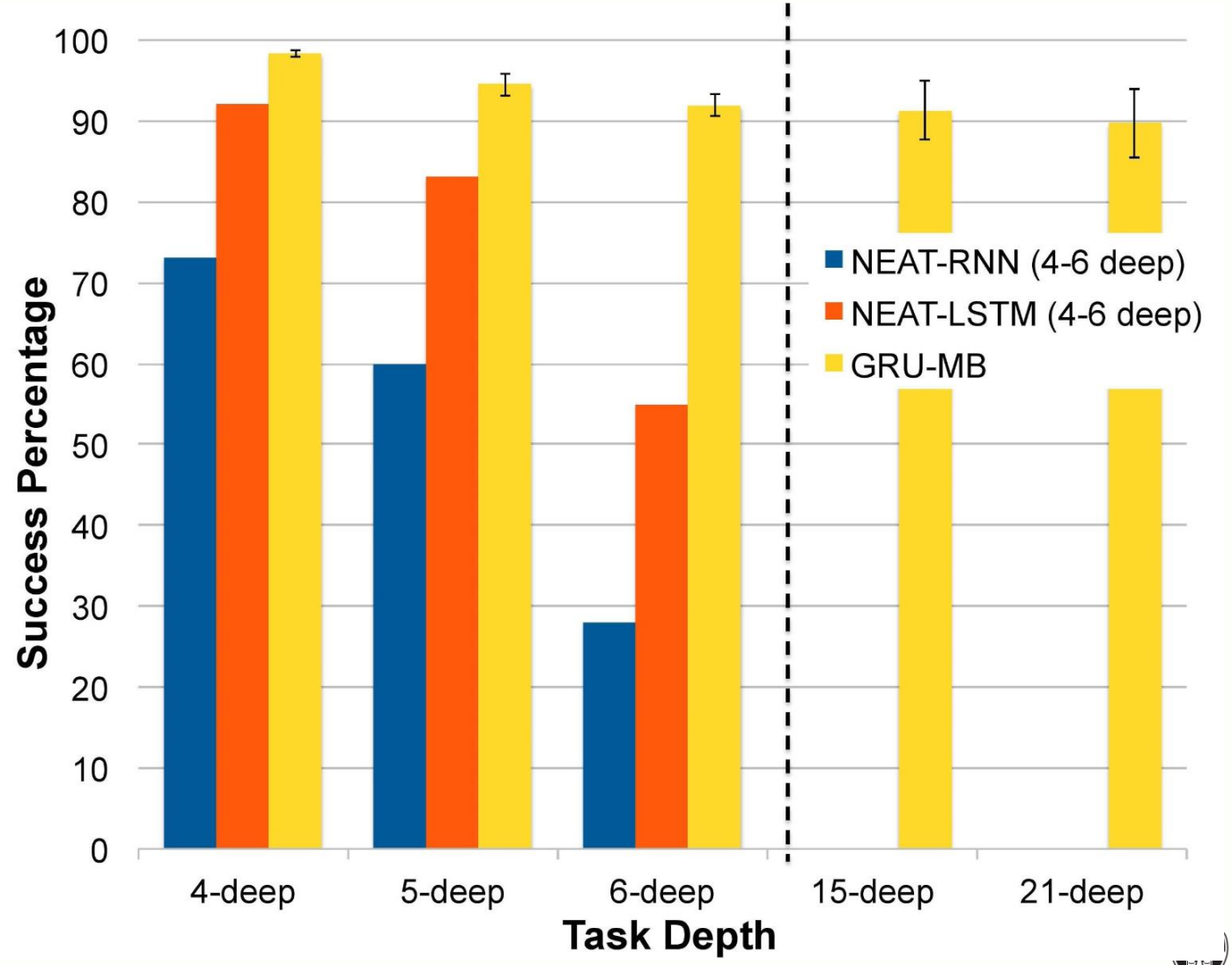


Recent GRU-MB results (GECCO 2017)



Oregon State University

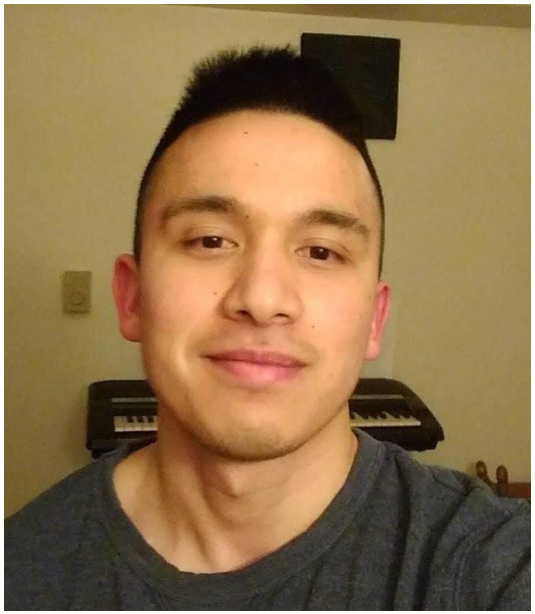
Recent GRU-MB results (GECCO 2017)



Oregon State University

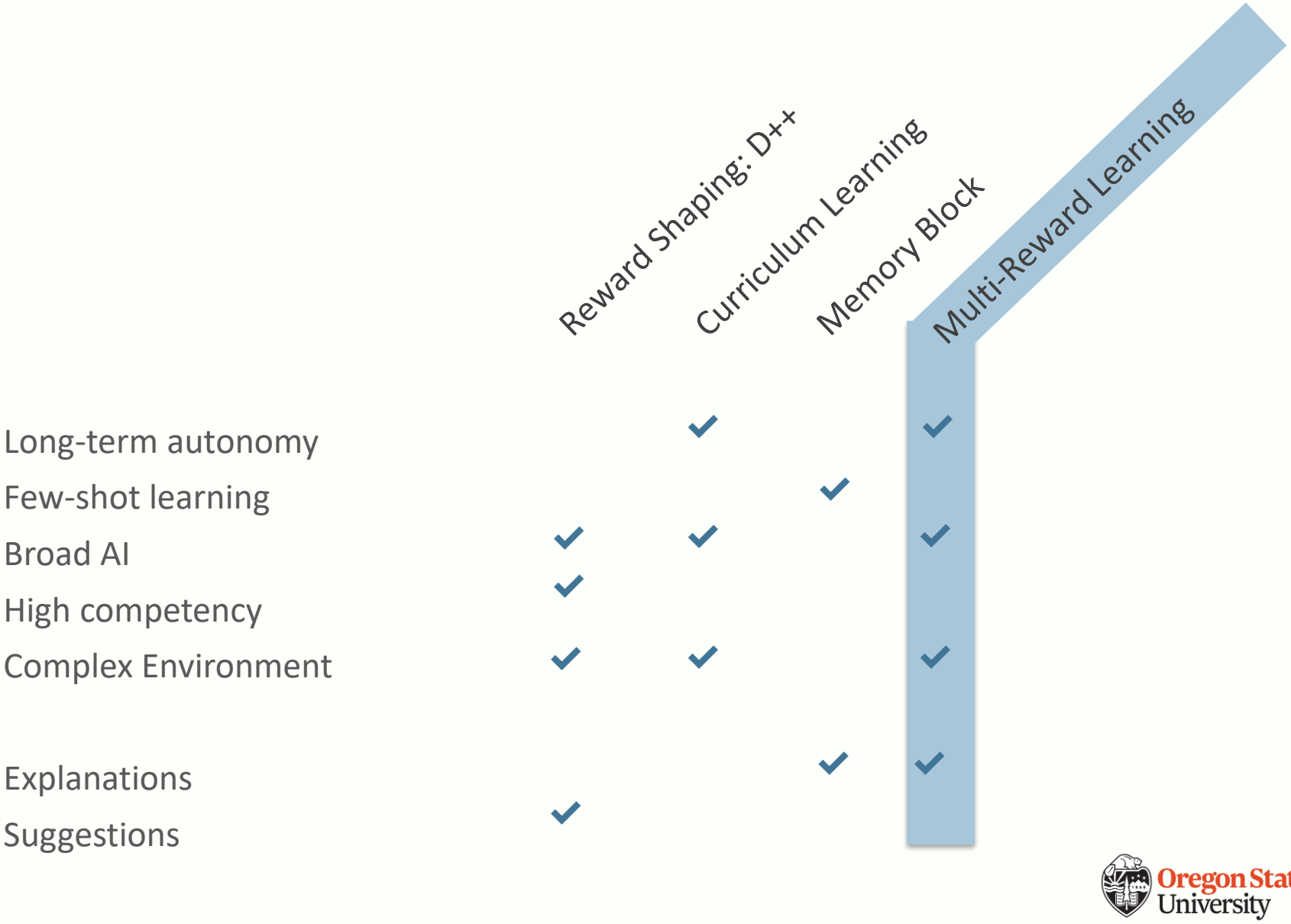
Key differences with LSTM ?

Ask Shaw:



Oregon State University

What will it take to get where we're going ?



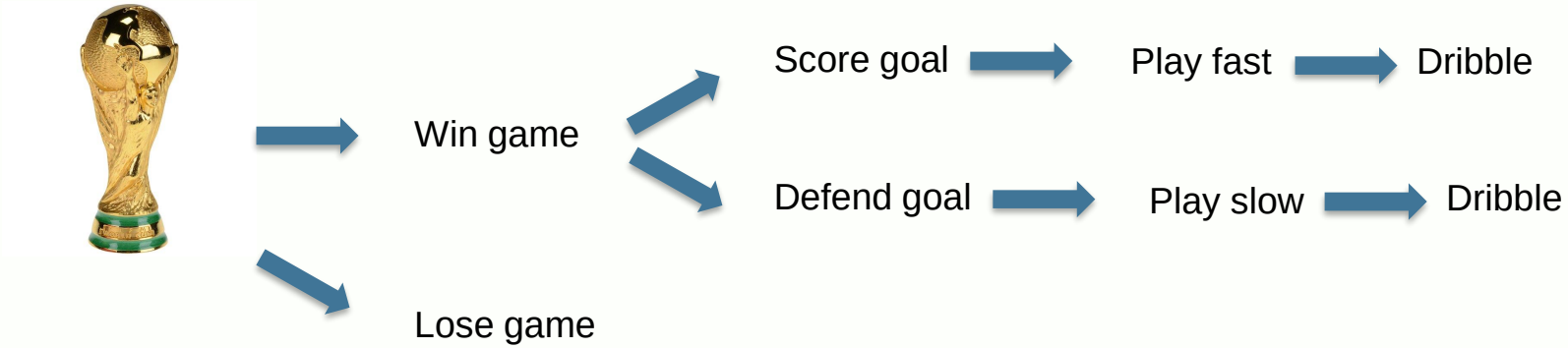
Oregon State University

So I want to learn how to play Soccer

Are these good rewards?

Oregon State University

Reward Alignment



What reward in what state at what time support “big” goal?

How do we decide reward structure?

Win / loss?

Do Experiments (#)

Survived (Y/N)

Area Explored (m^2)

Discovered Life (Y/N)

Oregon State University

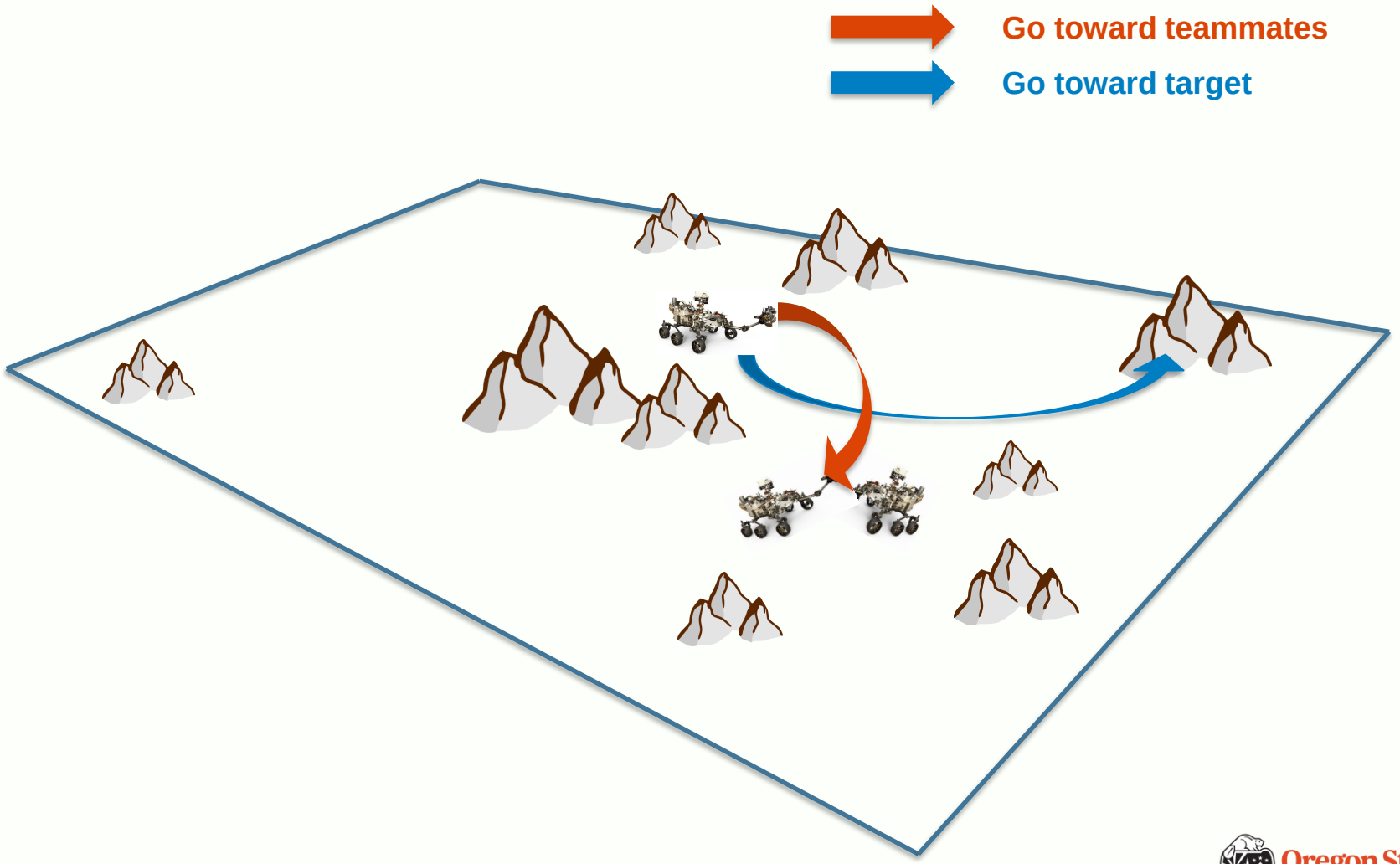
Multiple Reward Signals

The key to intelligent team behavior in complex systems is determining

“what matters when”

Oregon State University

Back to our Rover problem: 3 observations required



Oregon State University

What to do?

Measure alignment between high-level objective and immediate objectives

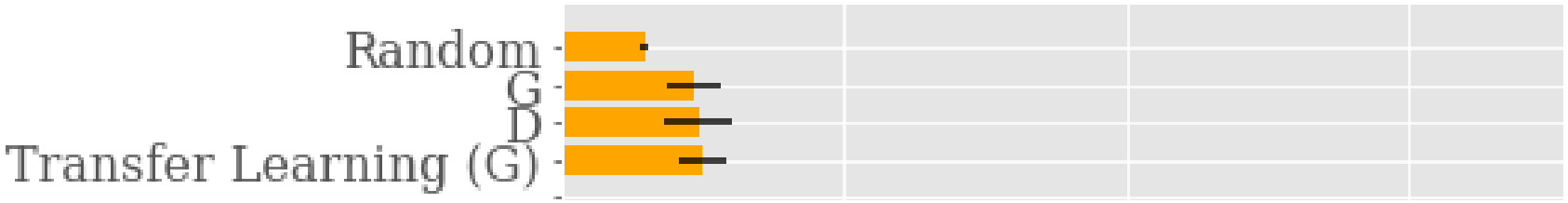
Pick immediate objective that supports high-level objective, here and now

How to calculate alignment?

Connor's talk later today!

Oregon State University

50 x 50 World with 30 Agents and 10 Targets



Average Reward with Converged Policies

Oregon State University

What will it take to get where we're going ?

Oregon State University

Key Takeaways

D++ : Stepping stone rewards

Curriculum learning

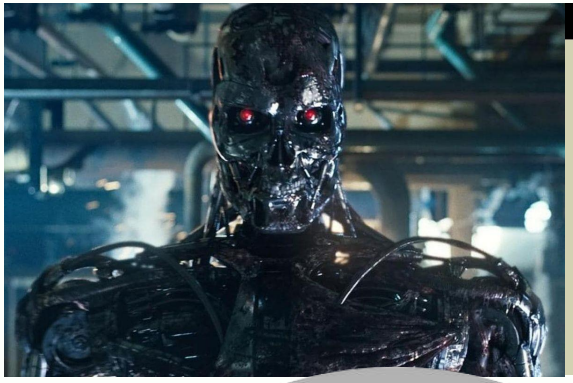
Memory Blocks

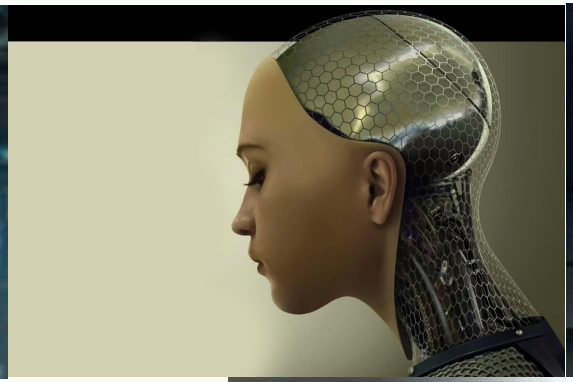
Multi-reward Learning

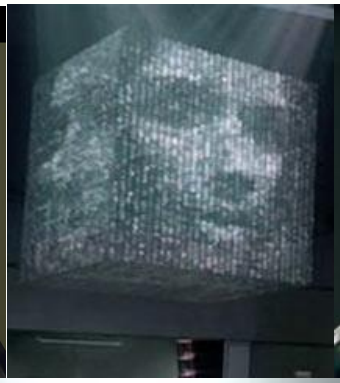
Suggestions

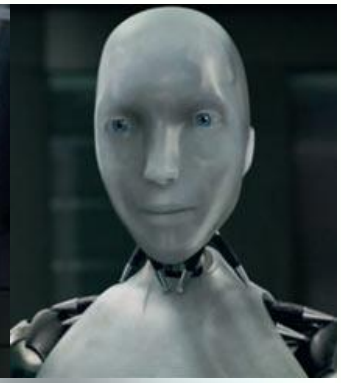
Oregon State University

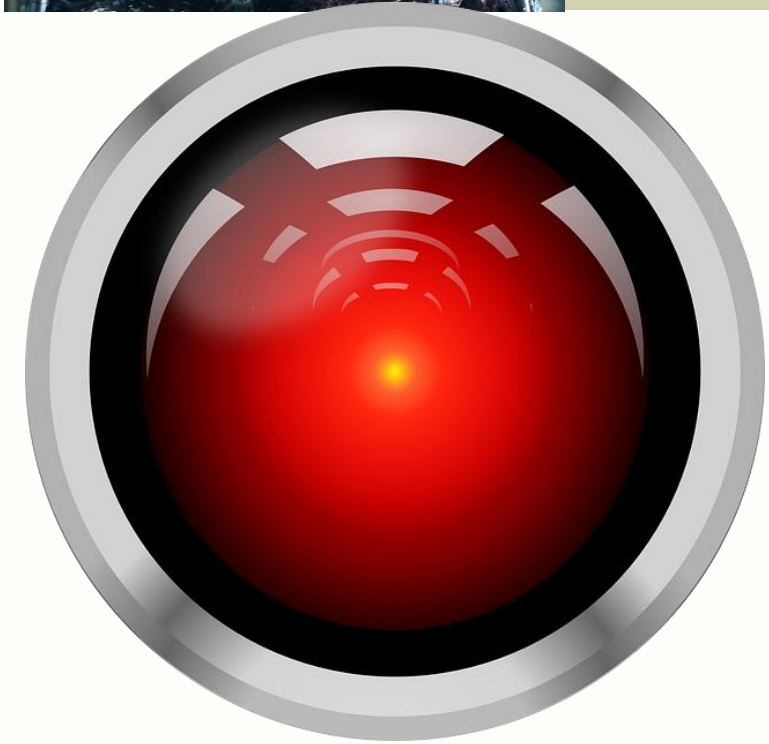
How is where we're going depicted ?







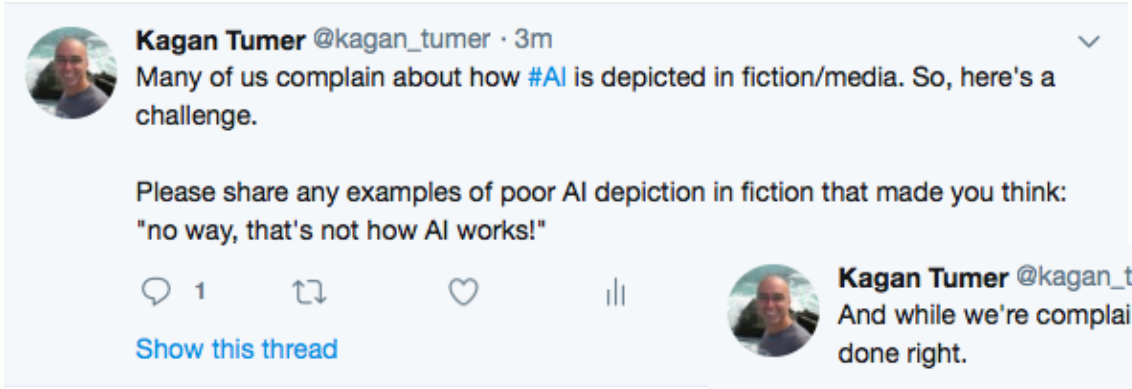


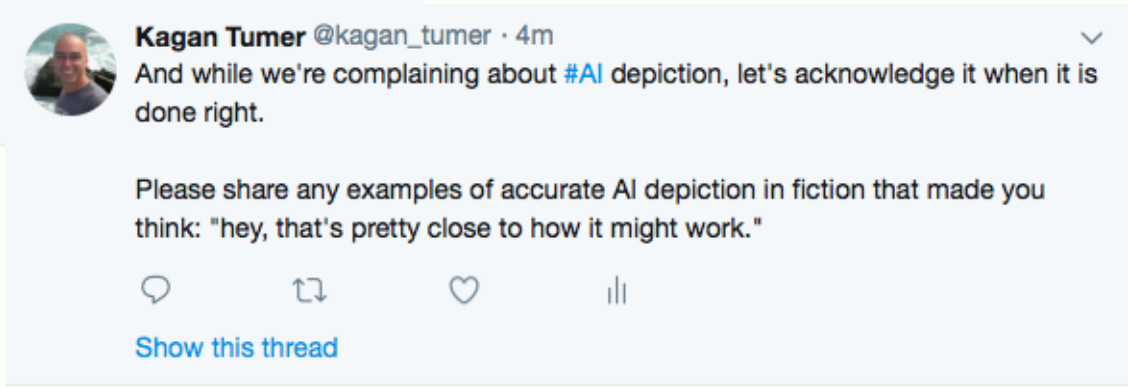


Oregon State University

Twitter Challenge

@kagan_tumer





Oregon State University

Autonomous Agents and Distributed Intelligence (AADI)

- Special thanks to:
  - Adrian Agogino
  - Sam Devlin

- AADI:
  - **Golden Rockefeller**
  - **Connor Yates**
  - **Shauharda Khadka**
  - Scott Chow
  - Mitch Colby
  - Jen Jen Chung
  - Logan Yliniemi
  - Carrie Rebhuhn
  - William Curran
  - Andrew Gabler
  - Sepide Kharaghani
  - Aida Rahmattalabi

- Eric Klinkhammer
- Atil Iscen
- Scott Proper
- Matt Knudson
- Stephen Sills
- Brent Valenti
- Jaime Junell
- Newsha Khani
- Adam Bell
- Ehsan Nasroullahi

Oregon State University

Autonomous Agents and Distributed Intelligence (AADI)

- Special thanks to our funders:

DAIMLER
DAIMLER TRUCKS NORTH AMERICA

EPRI
ELECTRIC POWER
RESEARCH INSTITUTE

Oregon State University

Questions?

@kagan_tumer

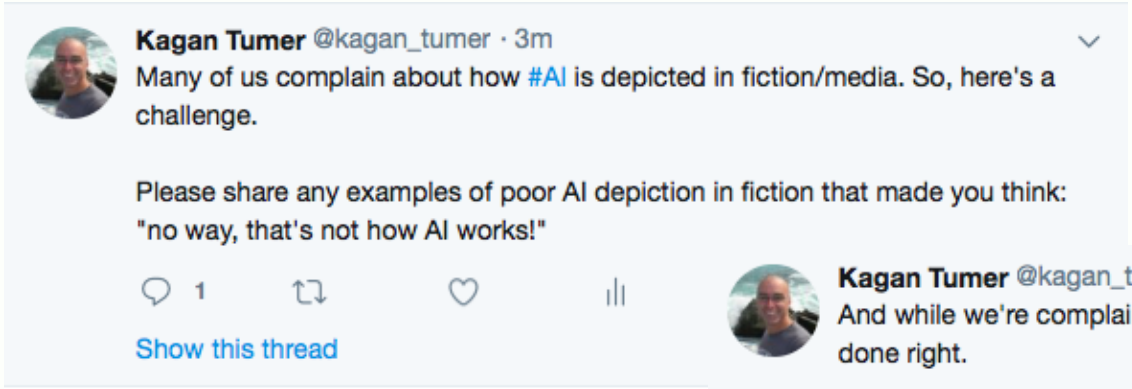
Kagan Tumer

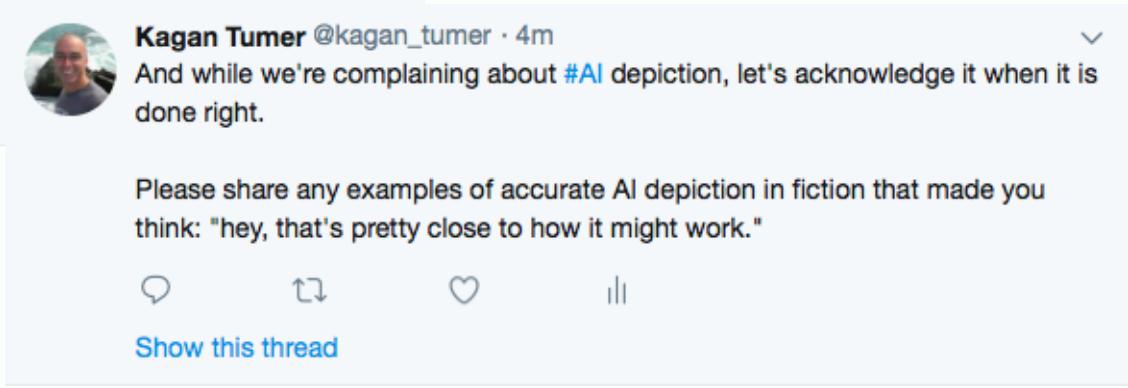
kagan.tumer@oregonstate.edu

Oregon State University

Twitter Challenge

@kagan_tumer





Oregon State University