Structured Computation and Representation in Deep Reinforcement Learning

Jessica B. Hamrick

DeepMind

ICLR Workshop on Deep Reinforcement
Learning Meets Structured Prediction

New Orleans, LA, USA
May 6, 2019

Understanding the World in Terms of Objects and Relations

Understanding the World in Terms of Objects and Relations

Understanding the World in Terms of Objects and Relations

Understanding the World in Terms of Objects and Relations

Structure

Structure

The arrangement of and ***relations*** between the ***parts*** or elements of something complex.

(Oxford English Dictionary)

Structure

The arrangement of and ***relations*** between the ***parts*** or elements of something complex.

(Oxford English Dictionary)

In humans: Evolution → Structure

Structure

The arrangement of and ***relations*** between the ***parts*** or elements of something complex.

(Oxford English Dictionary)

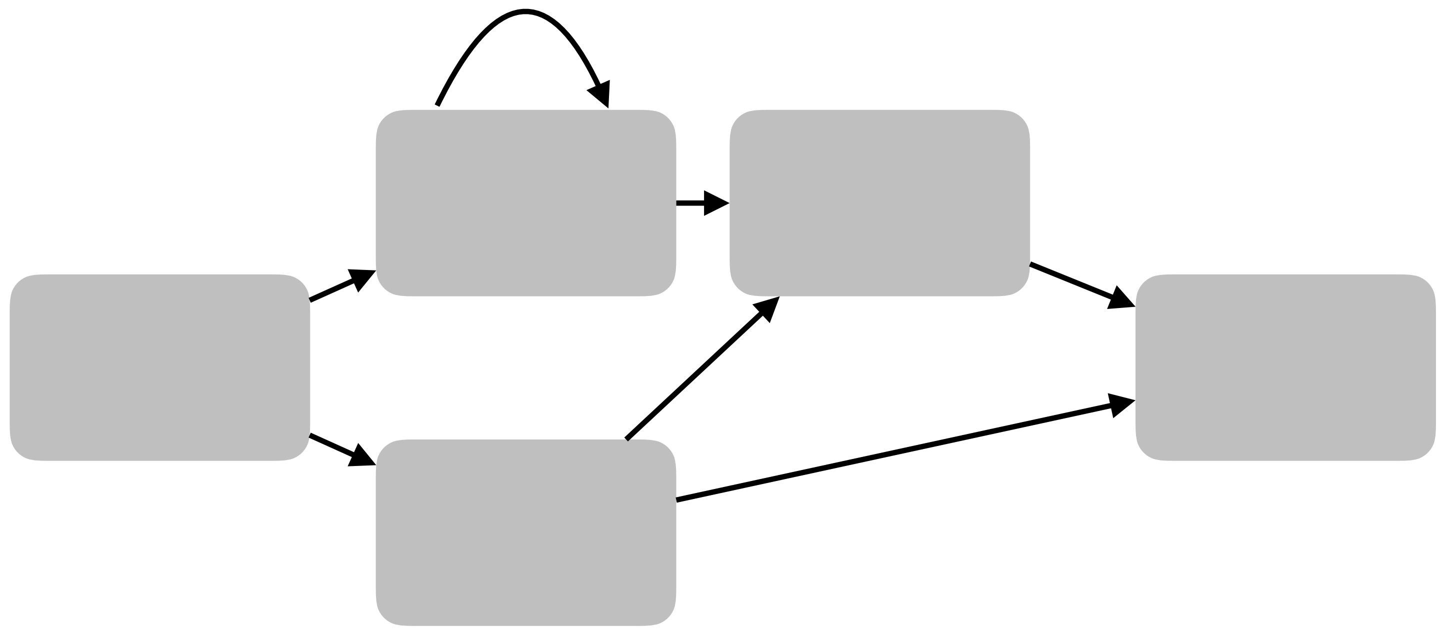
In humans: Evolution → Structure

In ML: Structure → Inductive bias / prior / regularizer

Two Types of Structure

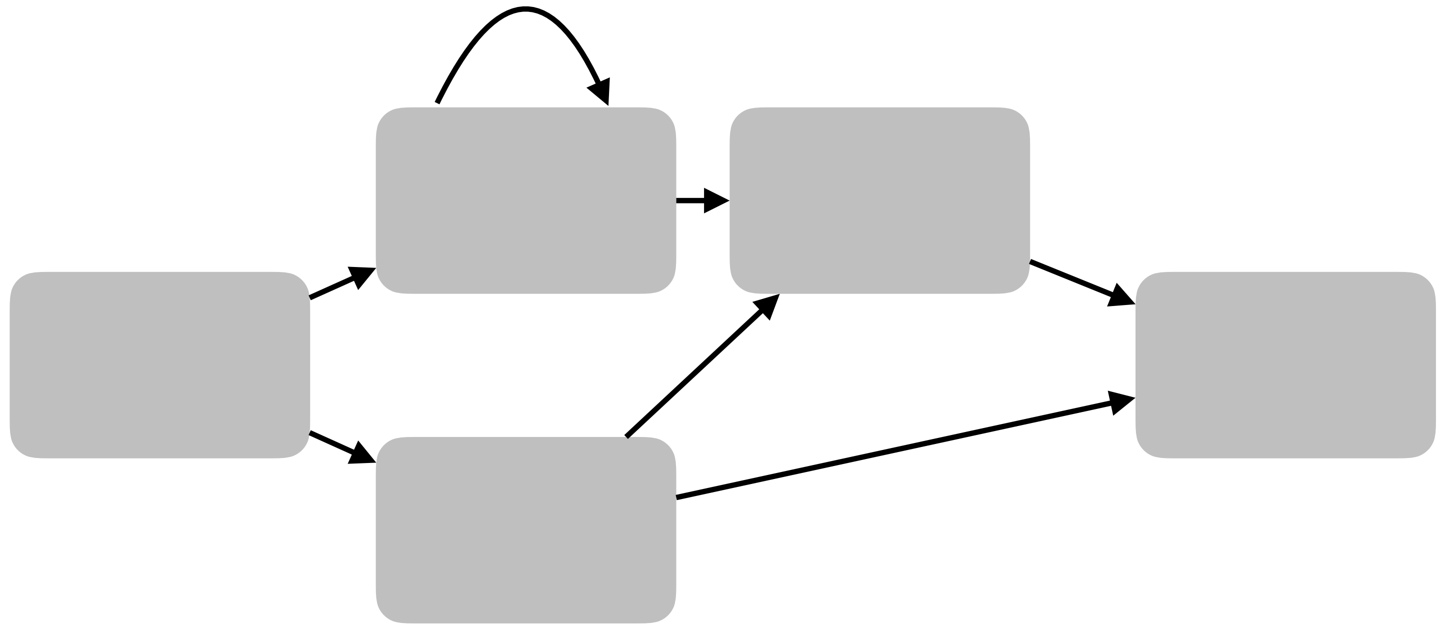
Two Types of Structure

1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

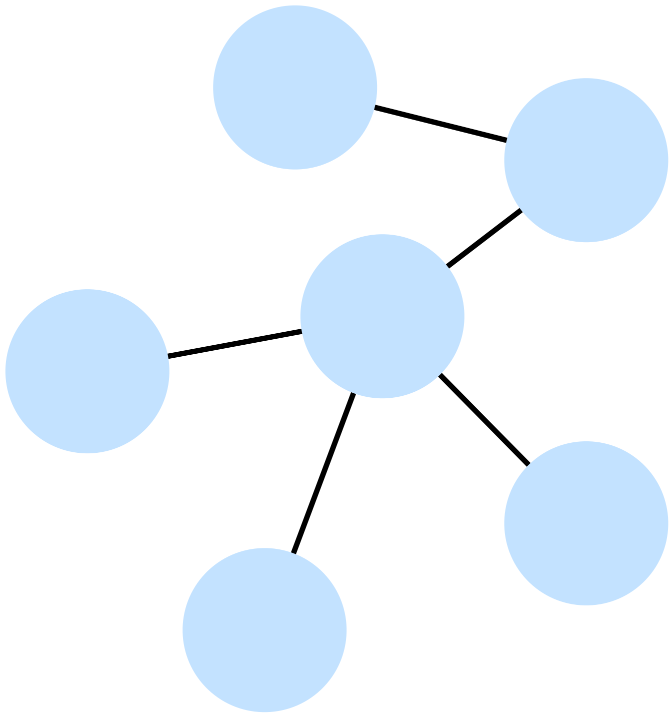


Two Types of Structure

1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.



2. **Structured Representation**: the format of the data that computations are performed over, e.g. sets, graphs, programs, etc.



A Caveat

A Caveat

Structured computation and structured representation go hand in hand!

A Caveat

Structured computation and structured representation go hand in hand!

Often, a structured representation entails
a particular structure in the computation, and vice versa.

A Caveat

Structured computation and structured representation go hand in hand!

Often, a structured representation entails
a particular structure in the computation, and vice versa.

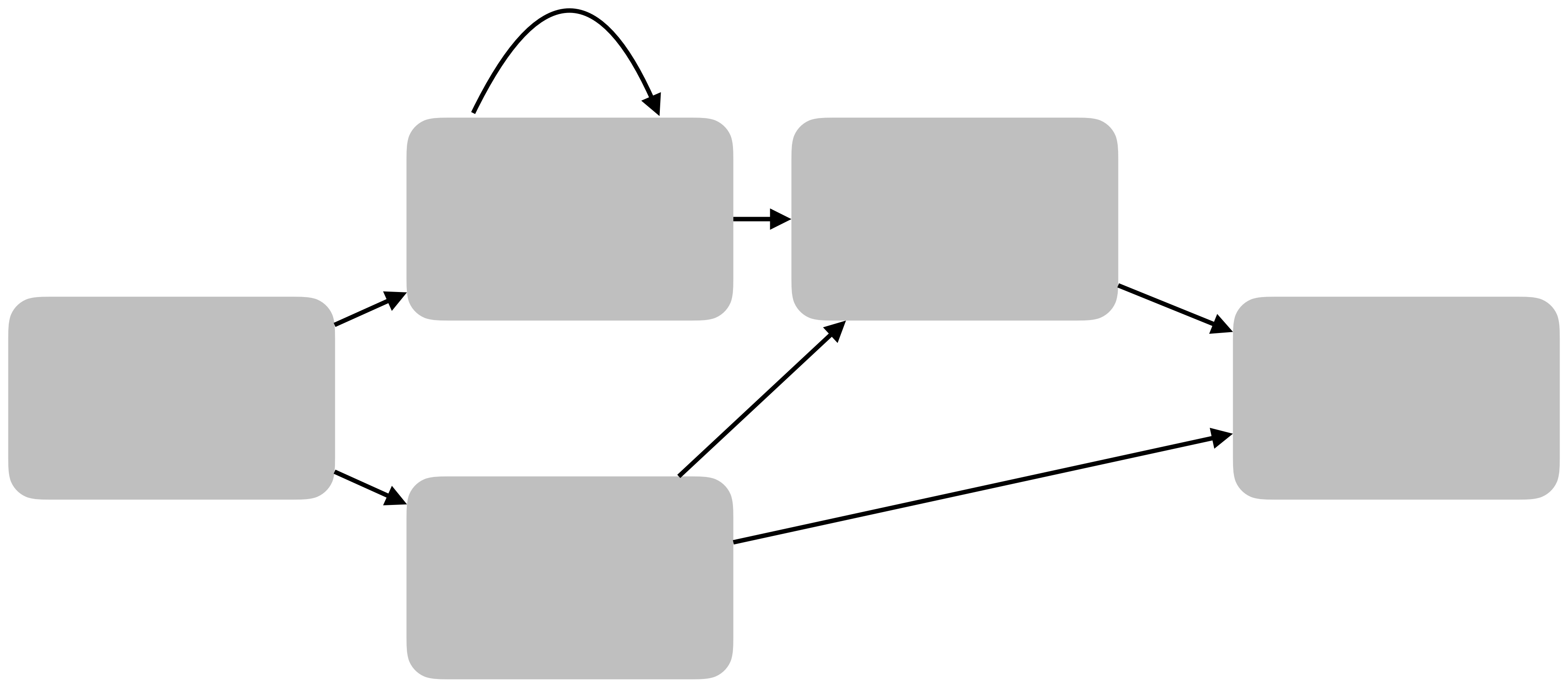
But it can still be useful to think separately about:
(1) how computations are arranged and composed, and
(2) the specific form of the representations.

A Caveat

Structured computation and structured representation go hand in hand!

Often, a structured representation entails
a particular structure in the computation, and vice versa.

But it can still be useful to think separately about:
(1) how computations are arranged and composed, and
(2) the specific form of the representations.

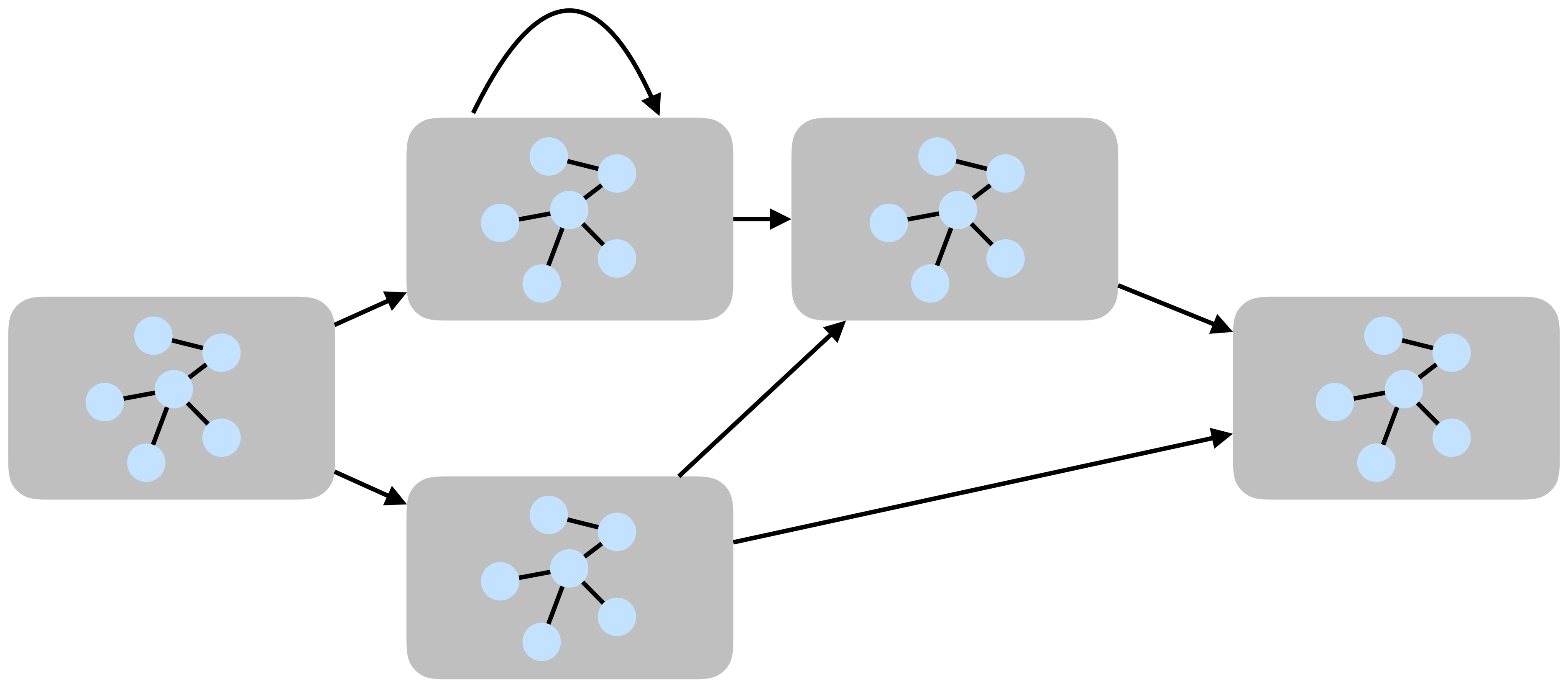


A Caveat

Structured computation and structured representation go hand in hand!

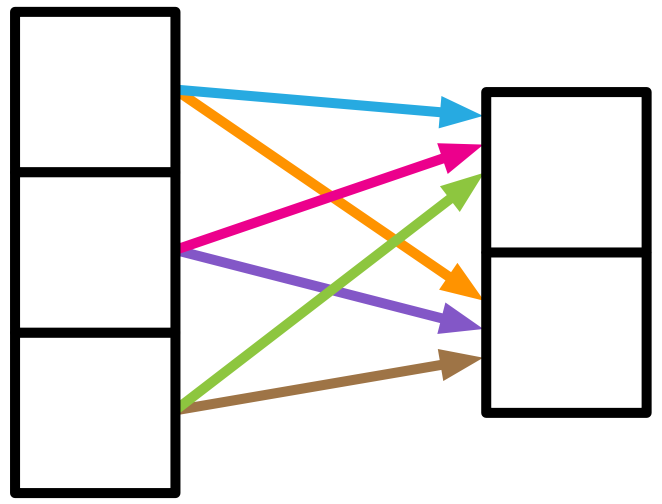
Often, a structured representation entails
a particular structure in the computation, and vice versa.

But it can still be useful to think separately about:
(1) how computations are arranged and composed, and
(2) the specific form of the representations.



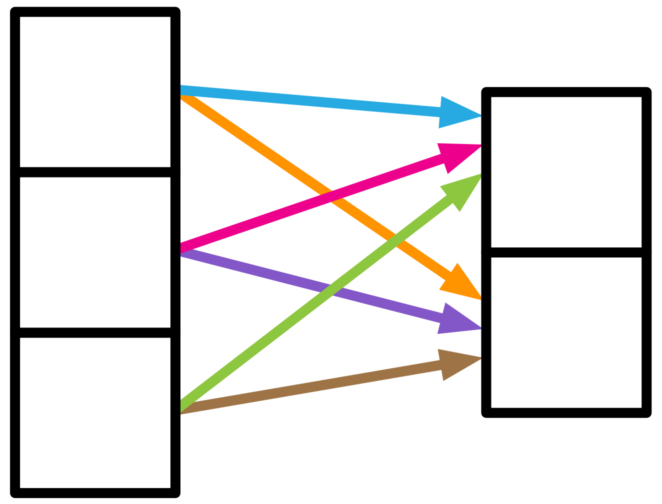
Structure in Deep Networks

Structure in Deep Networks

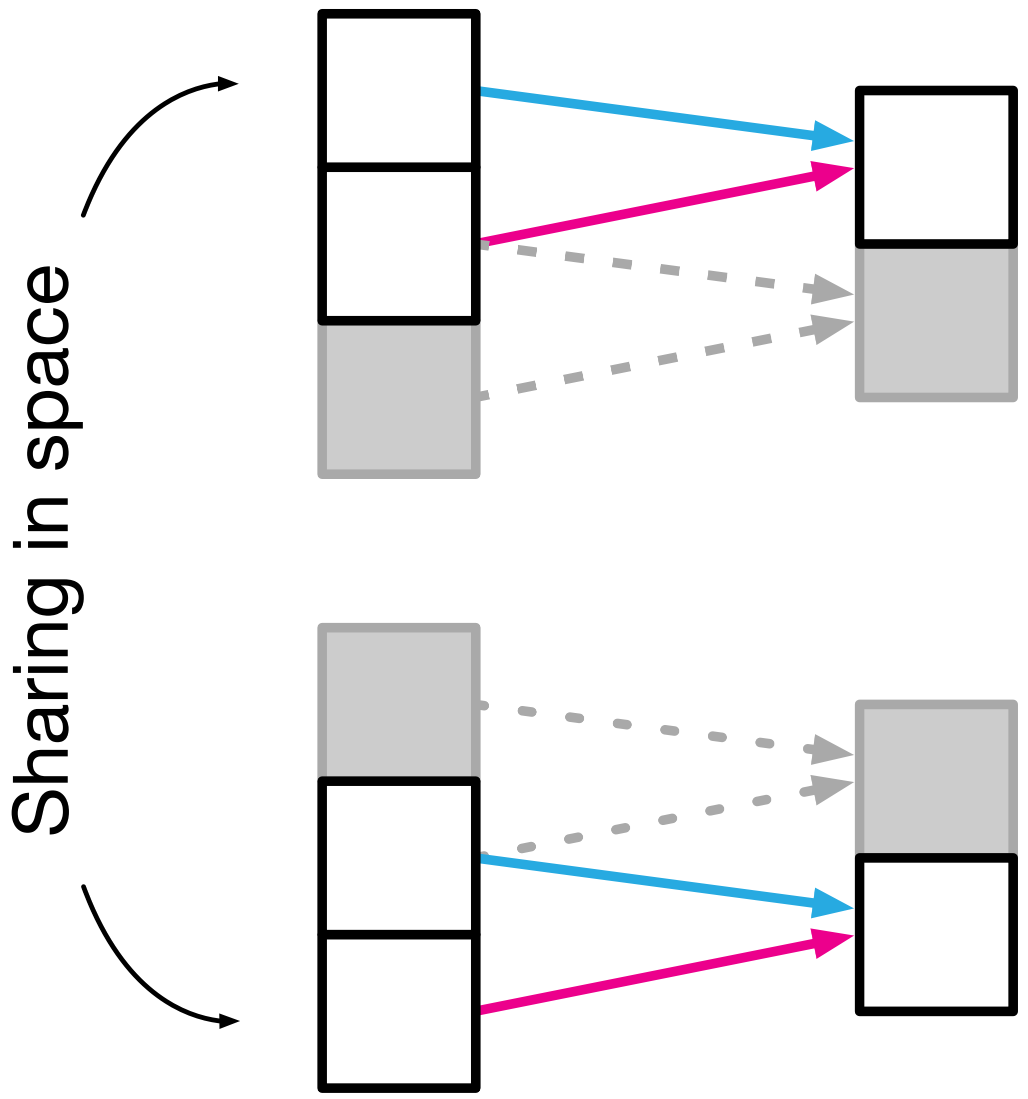


Fully-Connected Layer
Unstructured computation
Unstructured representation

Structure in Deep Networks

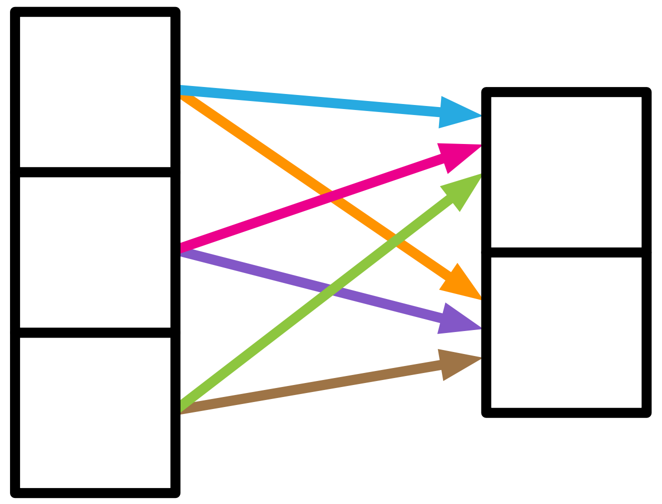


Fully-Connected Layer
Unstructured computation
Unstructured representation

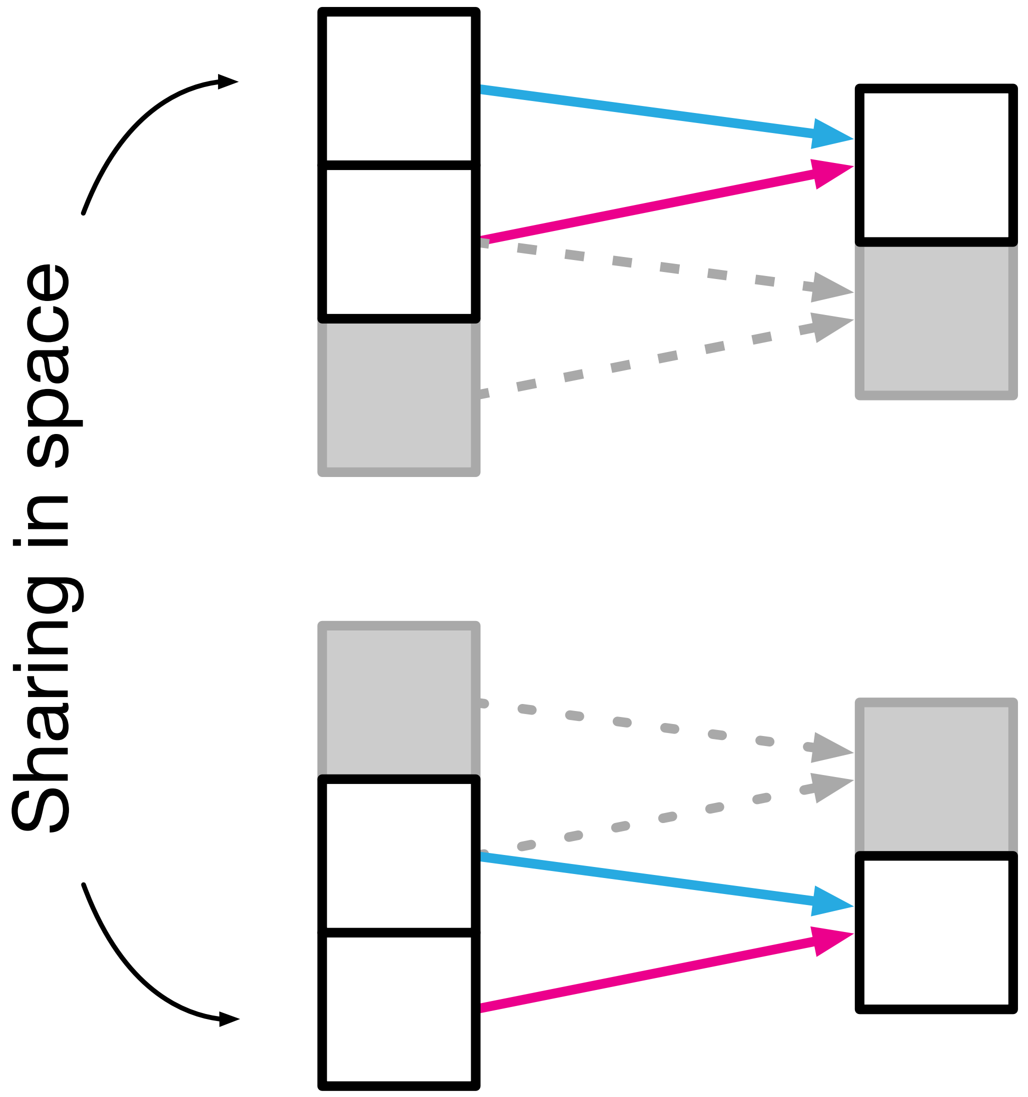


Convolutional Layer
Structured computation
Unstructured representation

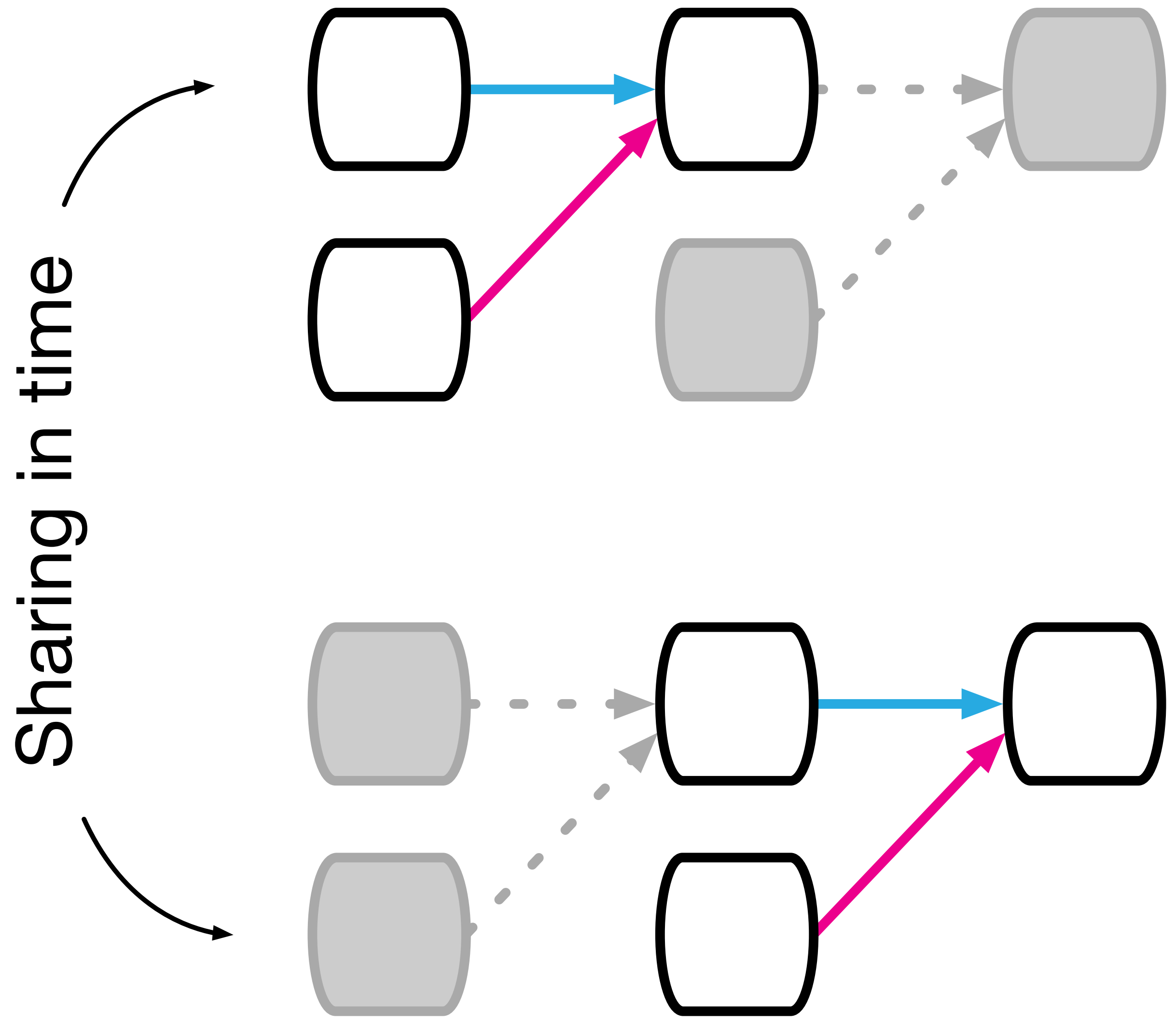
Structure in Deep Networks



Fully-Connected Layer
Unstructured computation
Unstructured representation

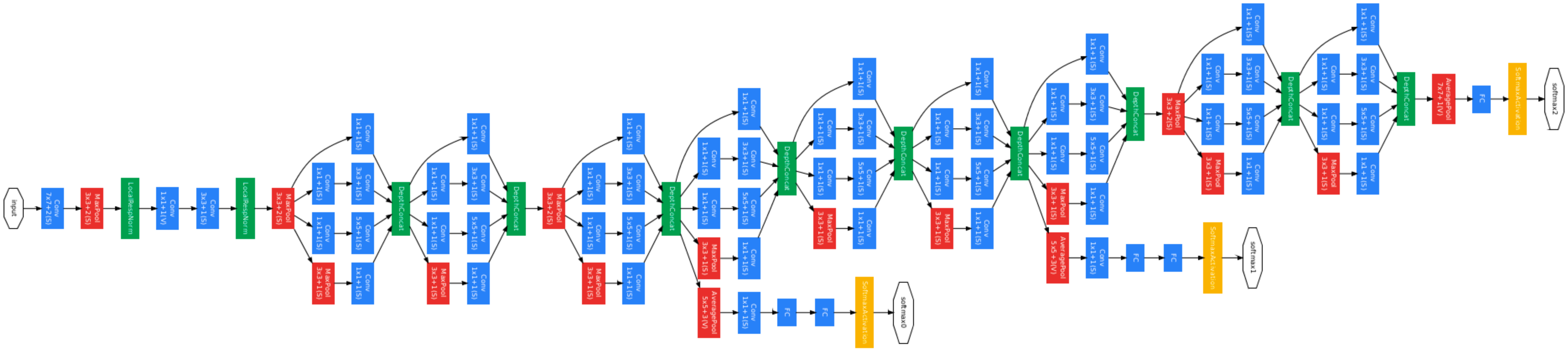


Convolutional Layer
Structured computation
Unstructured representation



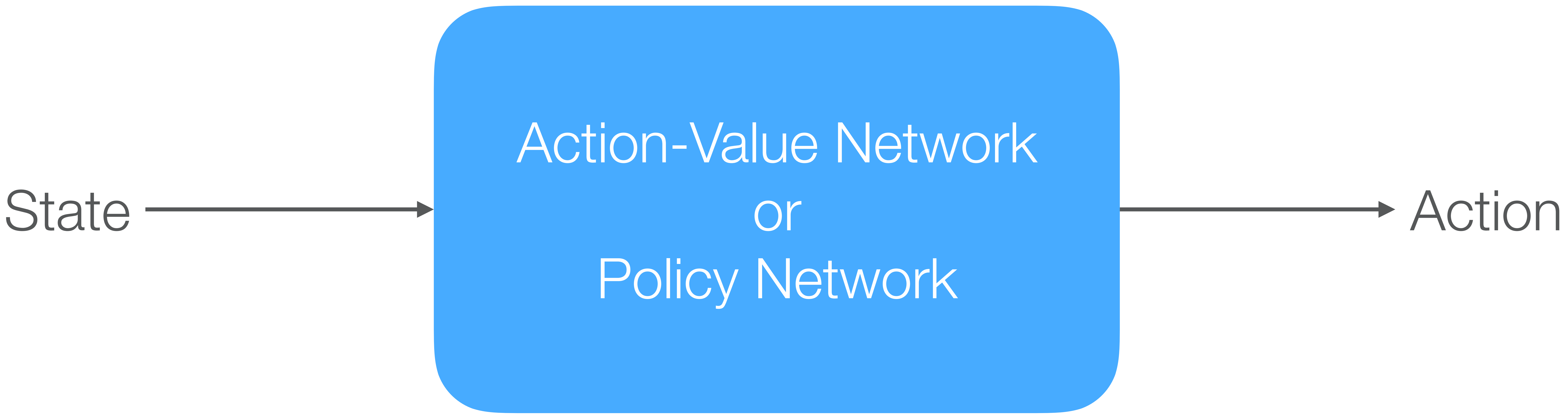
Recurrent Layer
Structured computation
Unstructured representation

Structure in Deep Networks



GoogLeNet with Inception modules, Szegedy et al. (2015)

Structure in Model-Free Deep RL

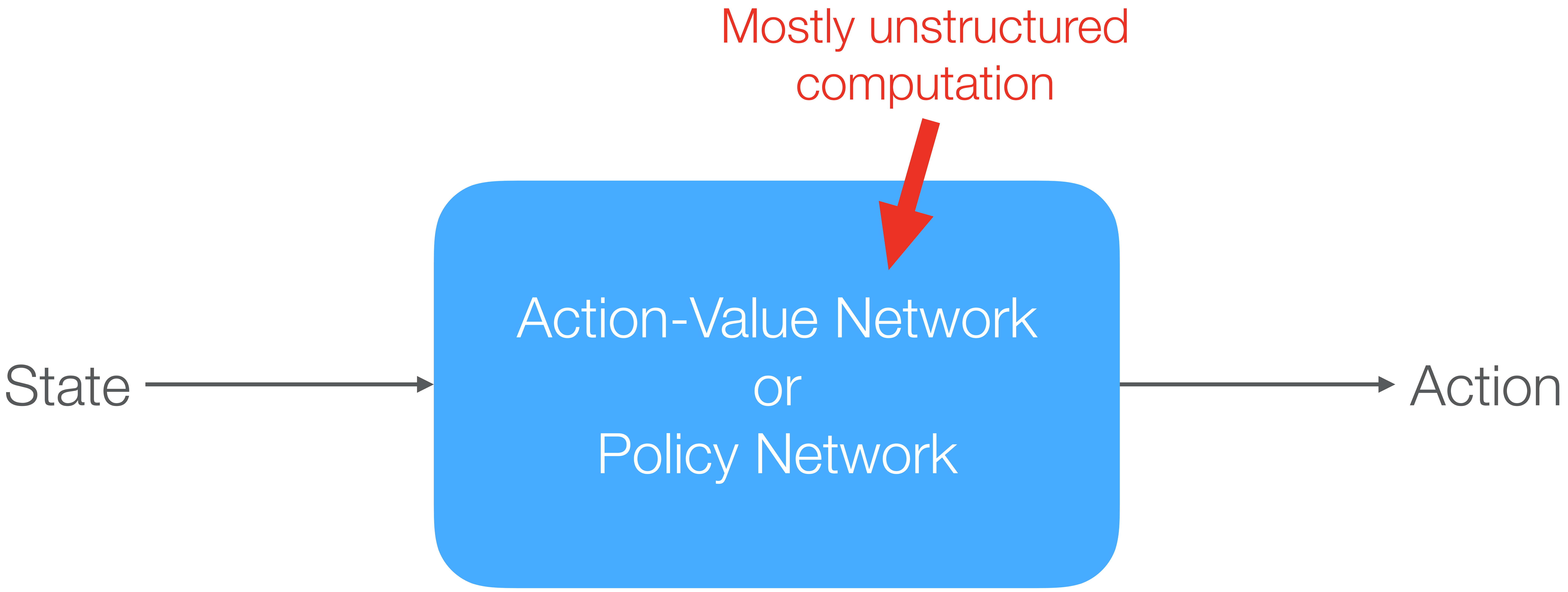


State

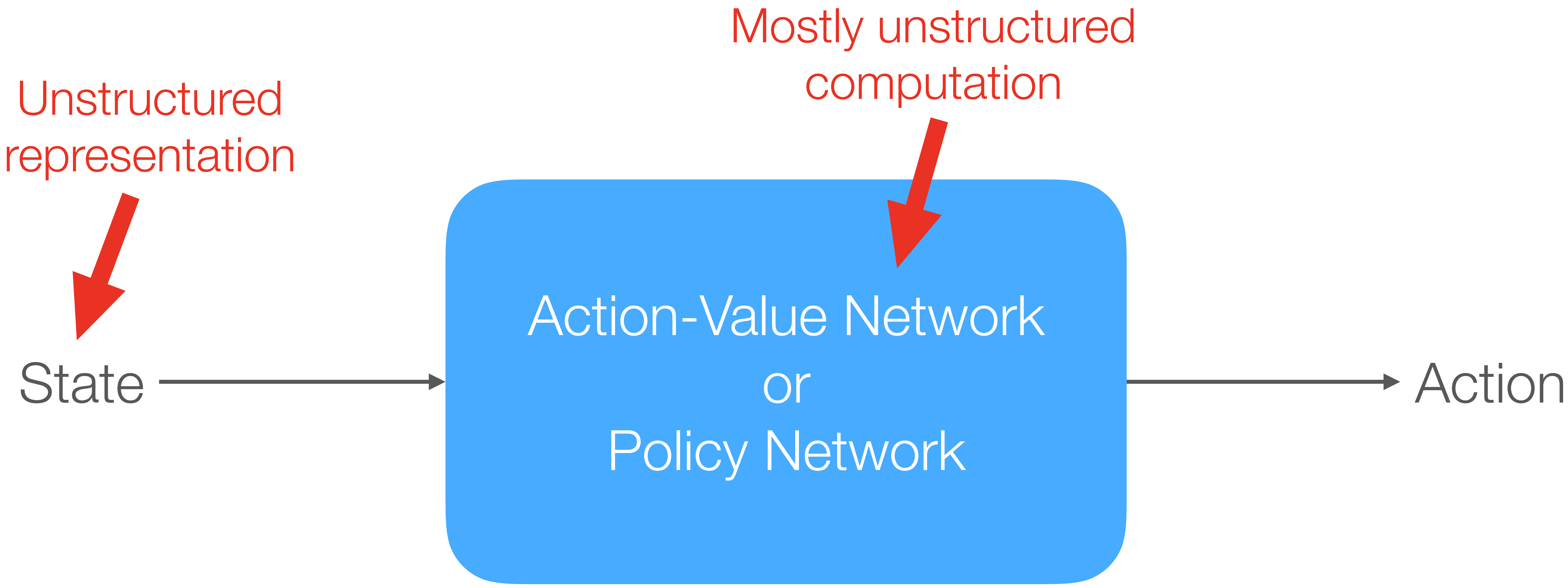
Action-Value Network
or
Policy Network

Action

Structure in Model-Free Deep RL



Structure in Model-Free Deep RL



Mostly unstructured computation

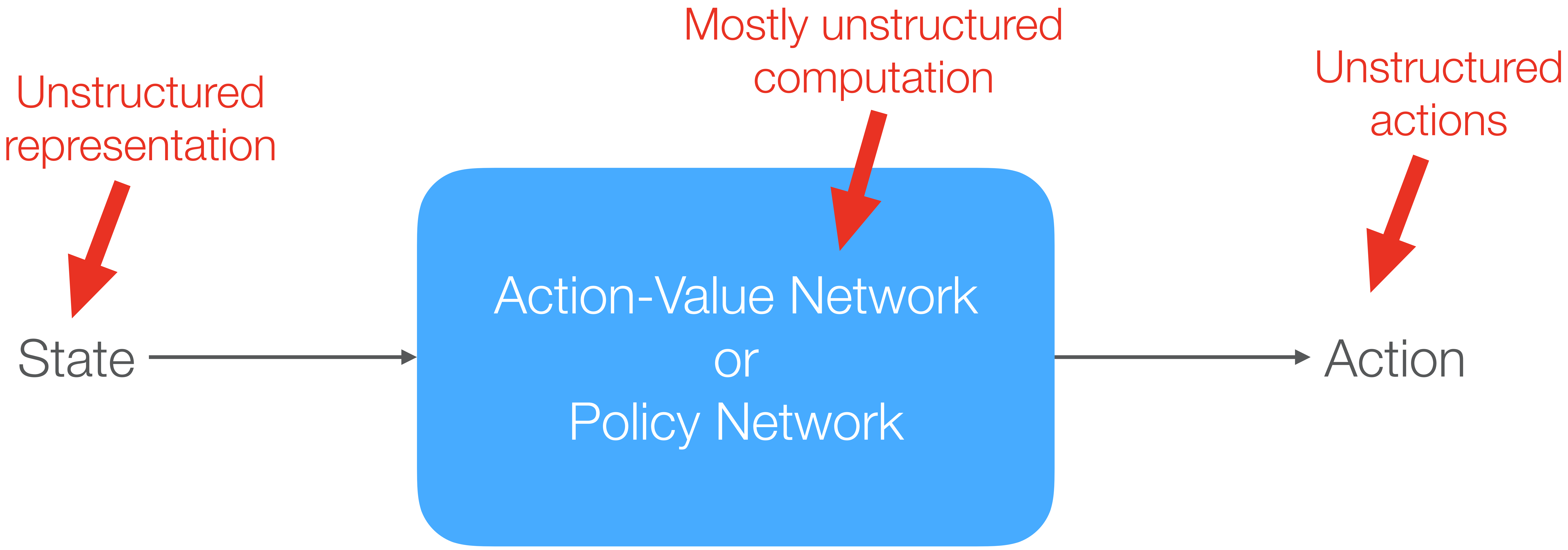
Unstructured representation

State

Action-Value Network
or
Policy Network

Action

Structure in Model-Free Deep RL



Unstructured representation

Mostly unstructured computation

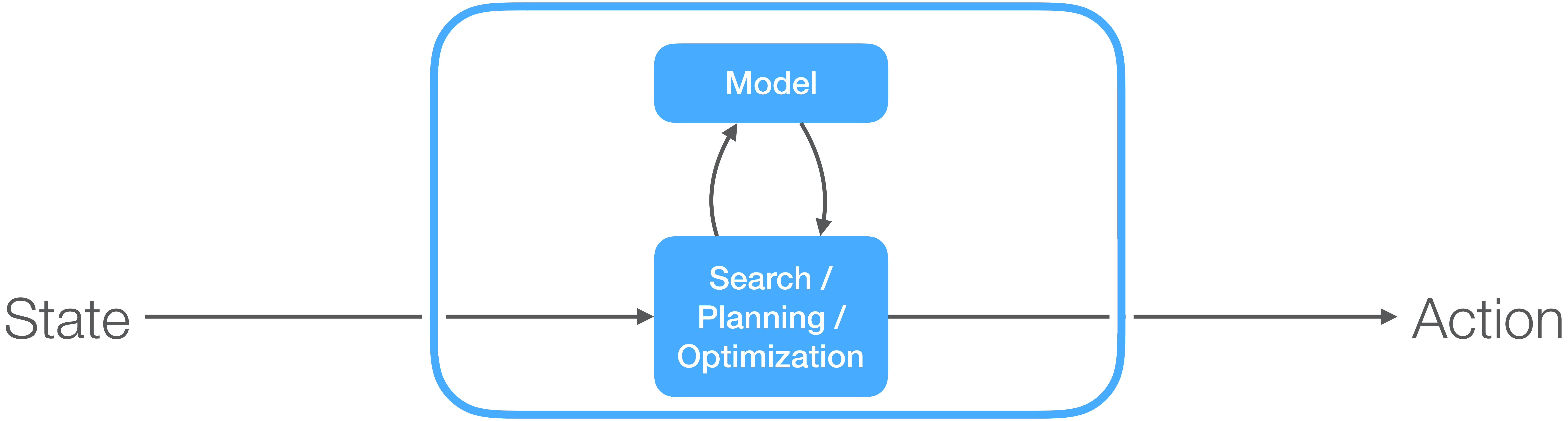
Unstructured actions

State

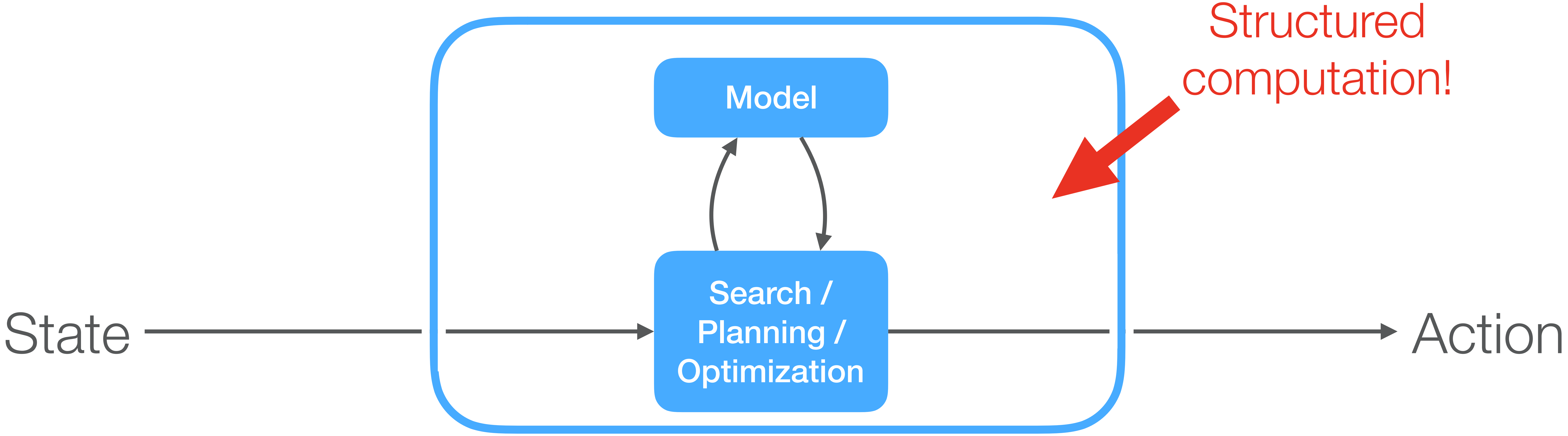
Action-Value Network or Policy Network

Action

Structure in Model-Based Deep RL



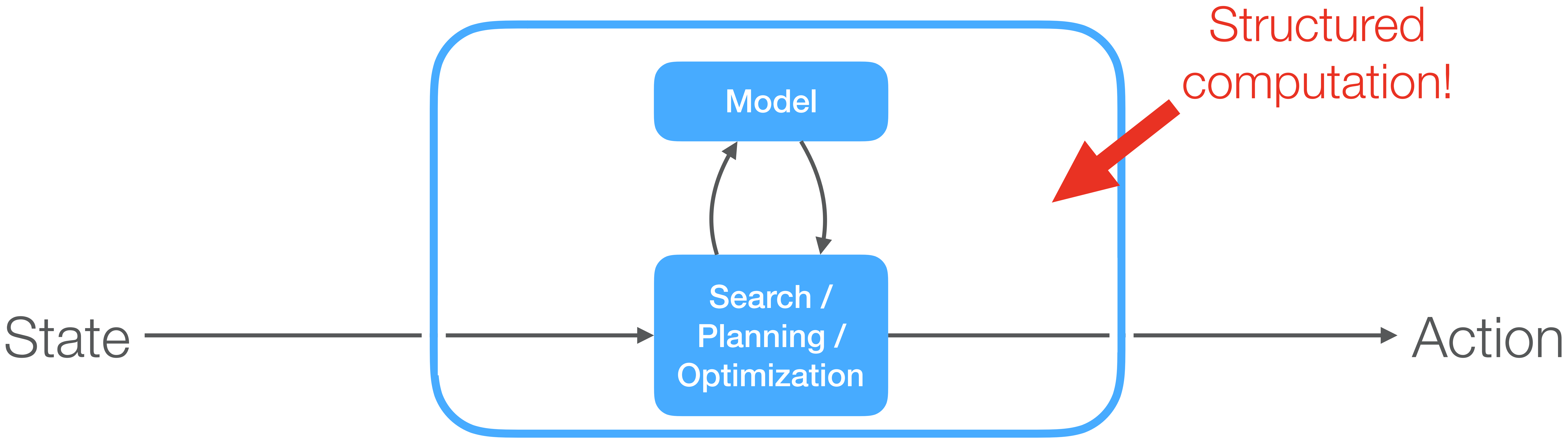
Structure in Model-Based Deep RL



Structured computation!

Structure in Model-Based Deep RL

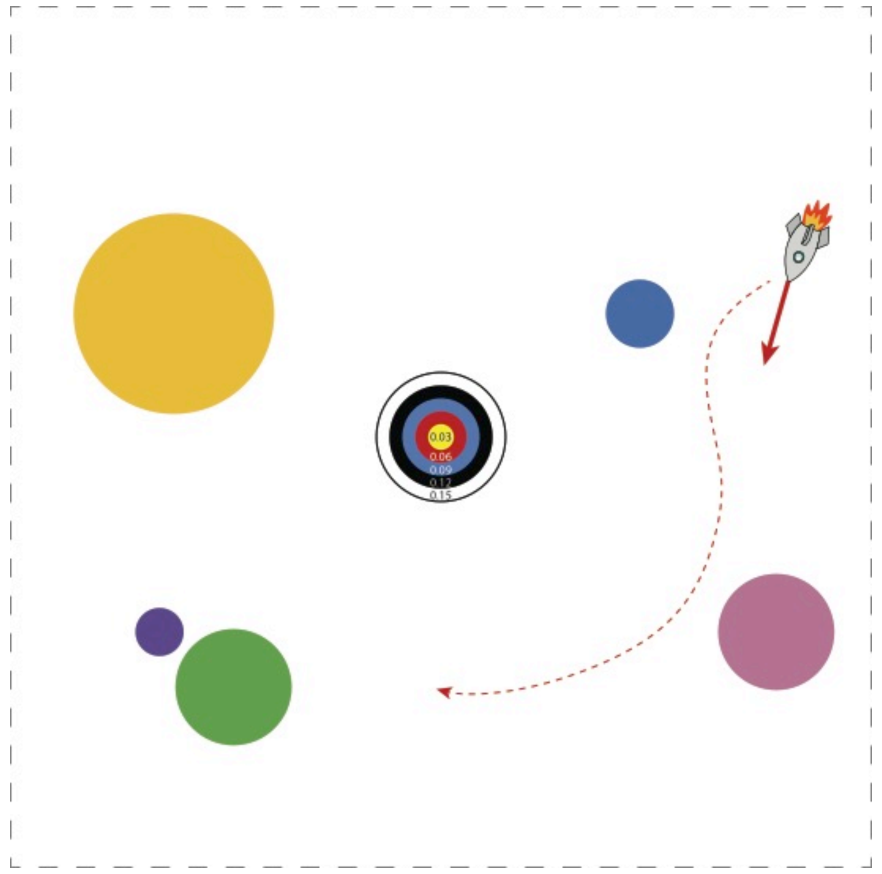
- AlphaGo (Silver et al., 2016)
- Imagination-Augmented Agents (Weber et al., 2017)
- Gradient Based Planning (Henaff et al., 2017)
- Value Prediction Networks (Oh et al., 2017)
- Universal Planning Networks (Srinivas et al., 2018)
- … and more! **See Hamrick (2019, *Current Opinion in Behavioral Sciences*) for a review.**



Outline

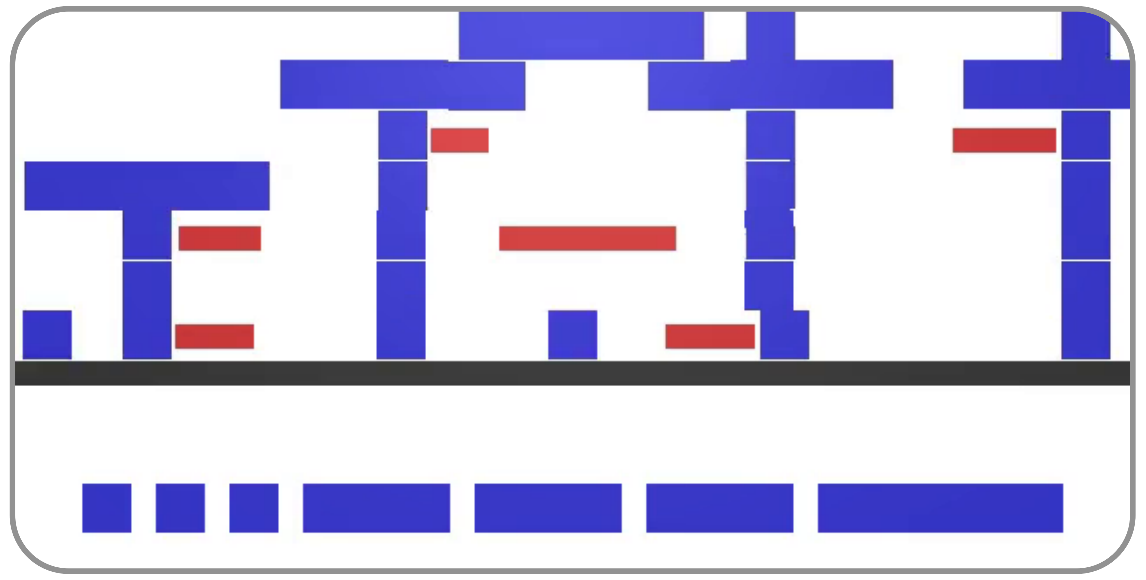
1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

   Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (2017, ICLR)



2. **Structured Representation**: the format of the data that computations are performed over.

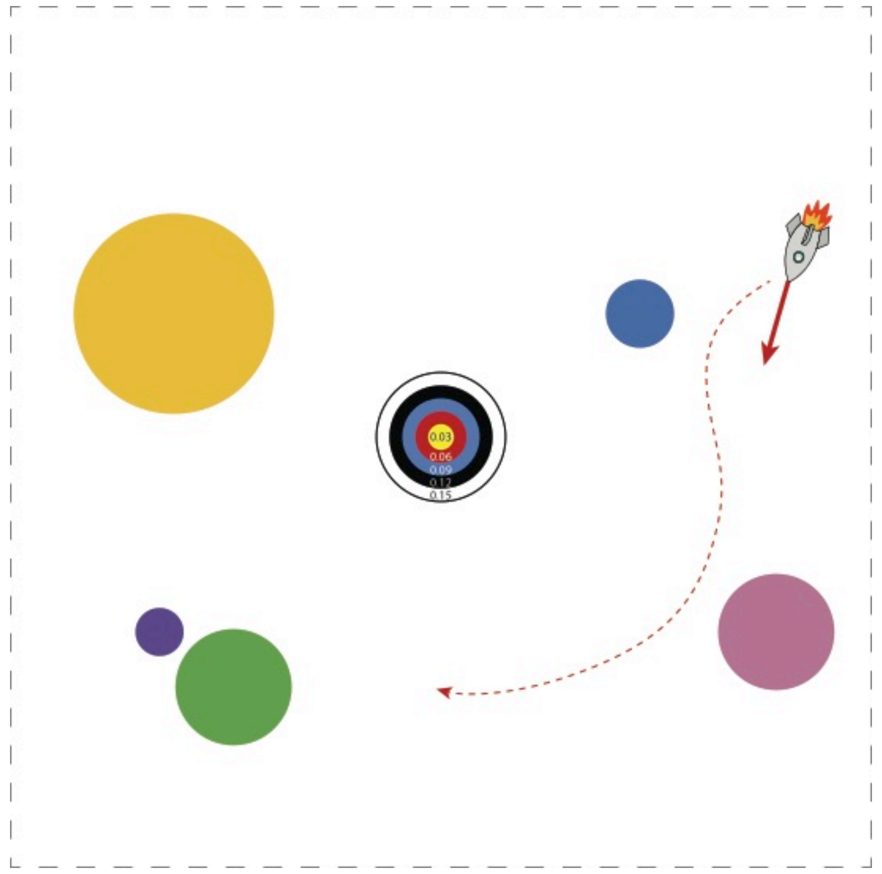
   Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018, arXiv)
   Bapst, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia, & Hamrick (2019, ICML)*



Outline

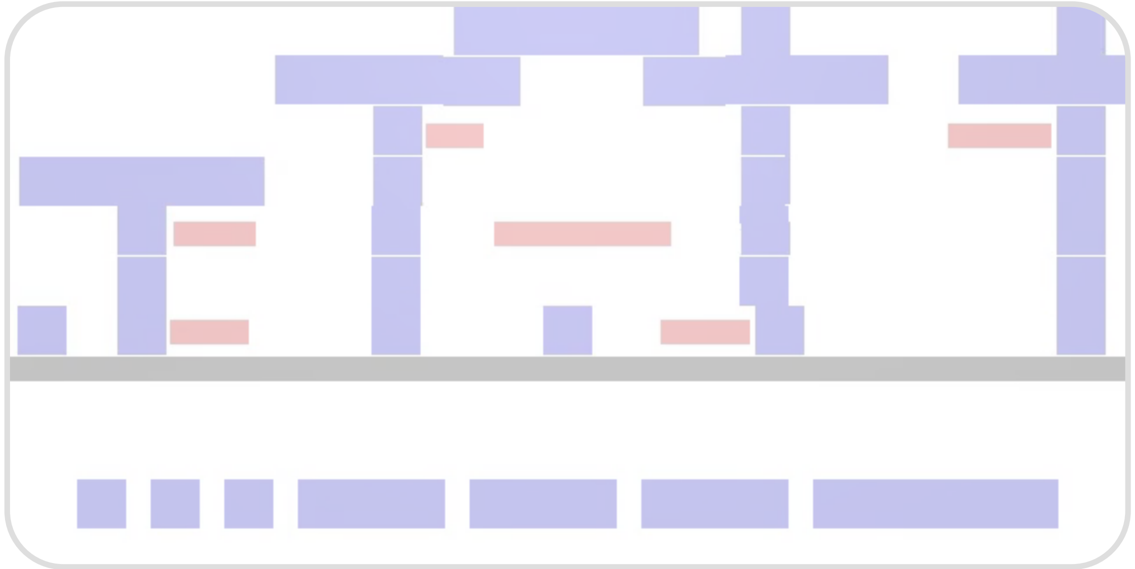
1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

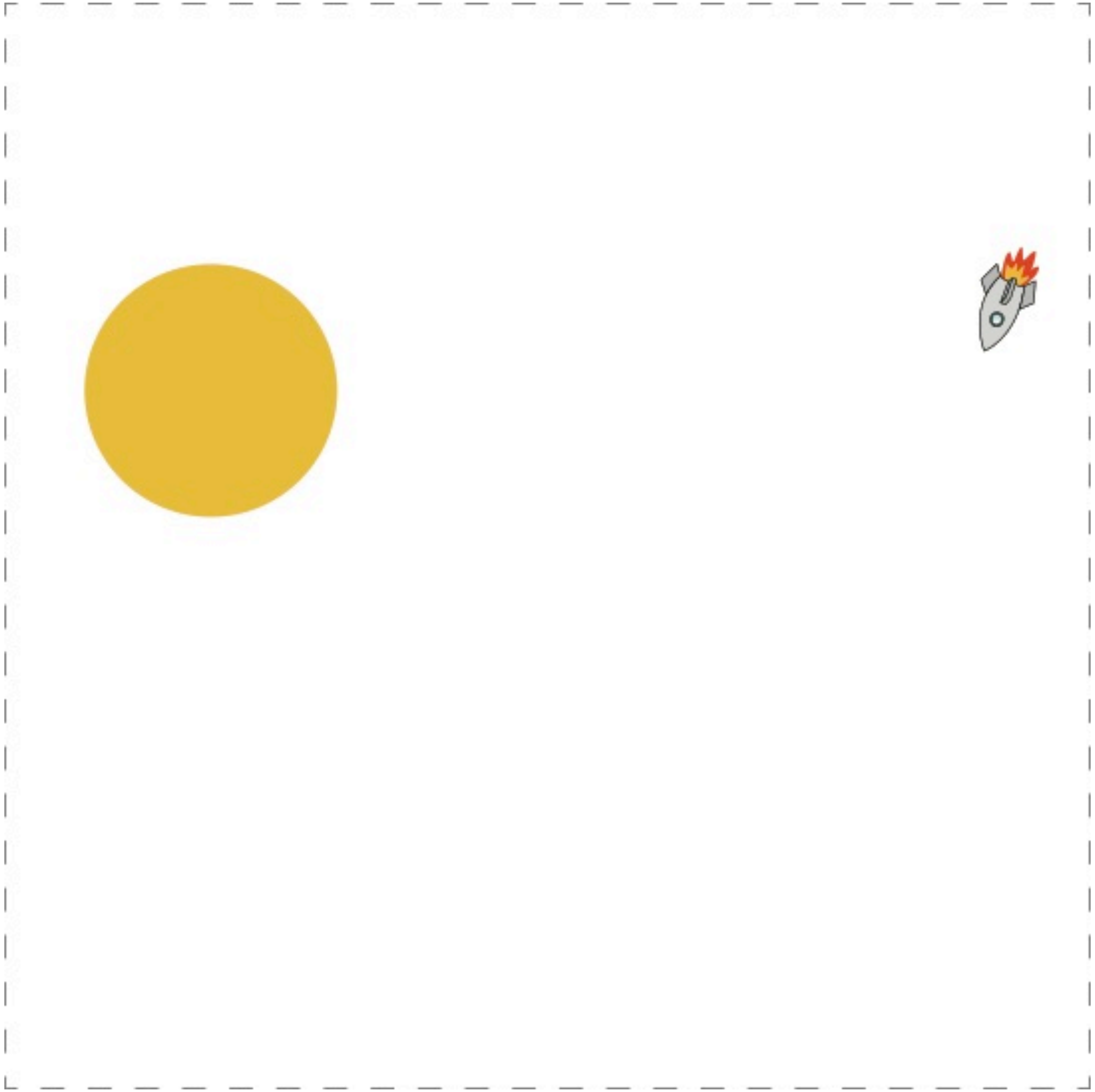
   Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (2017, ICLR)

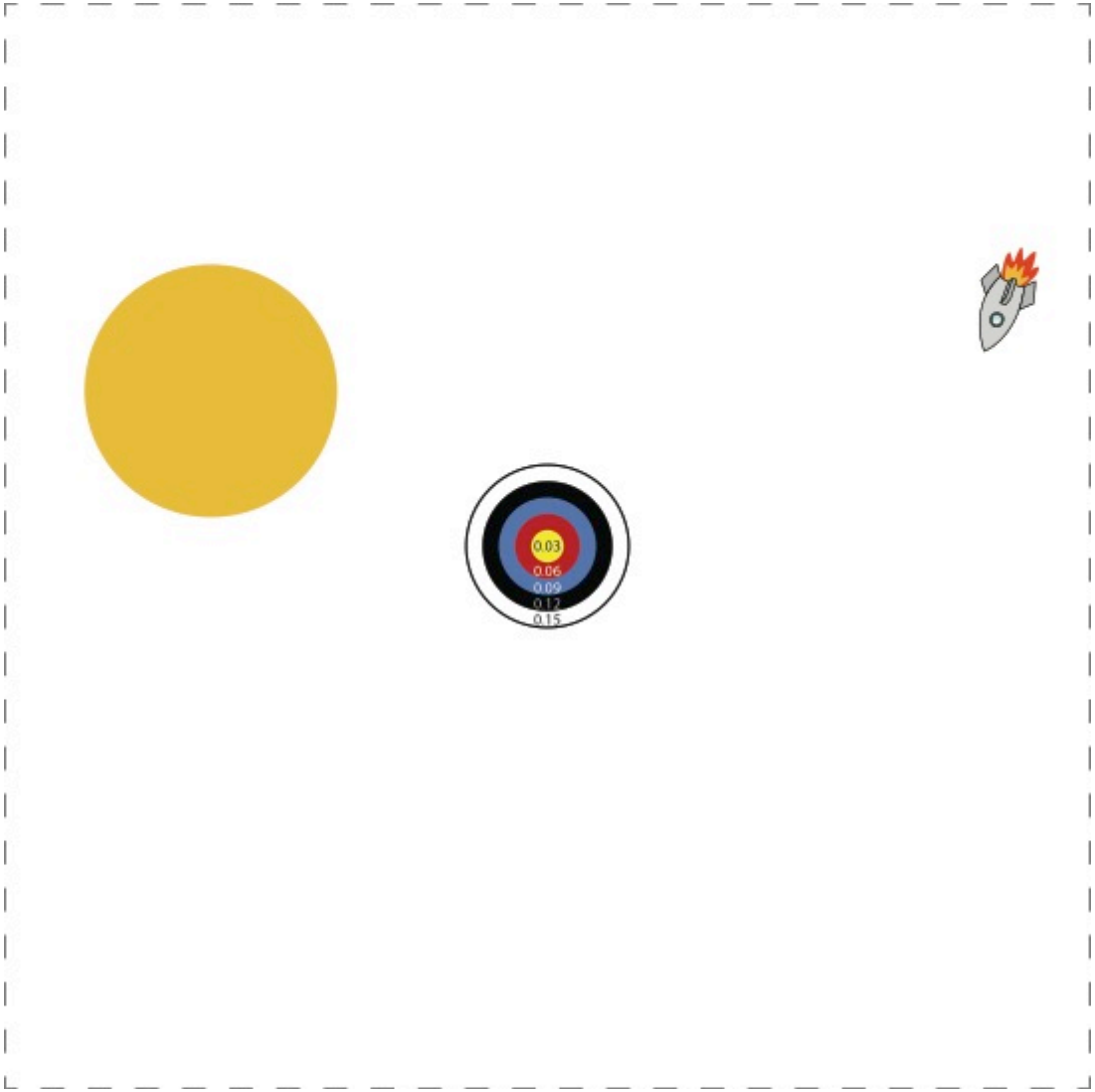


2. **Structured Representation**: the format of the data that computations are performed over.

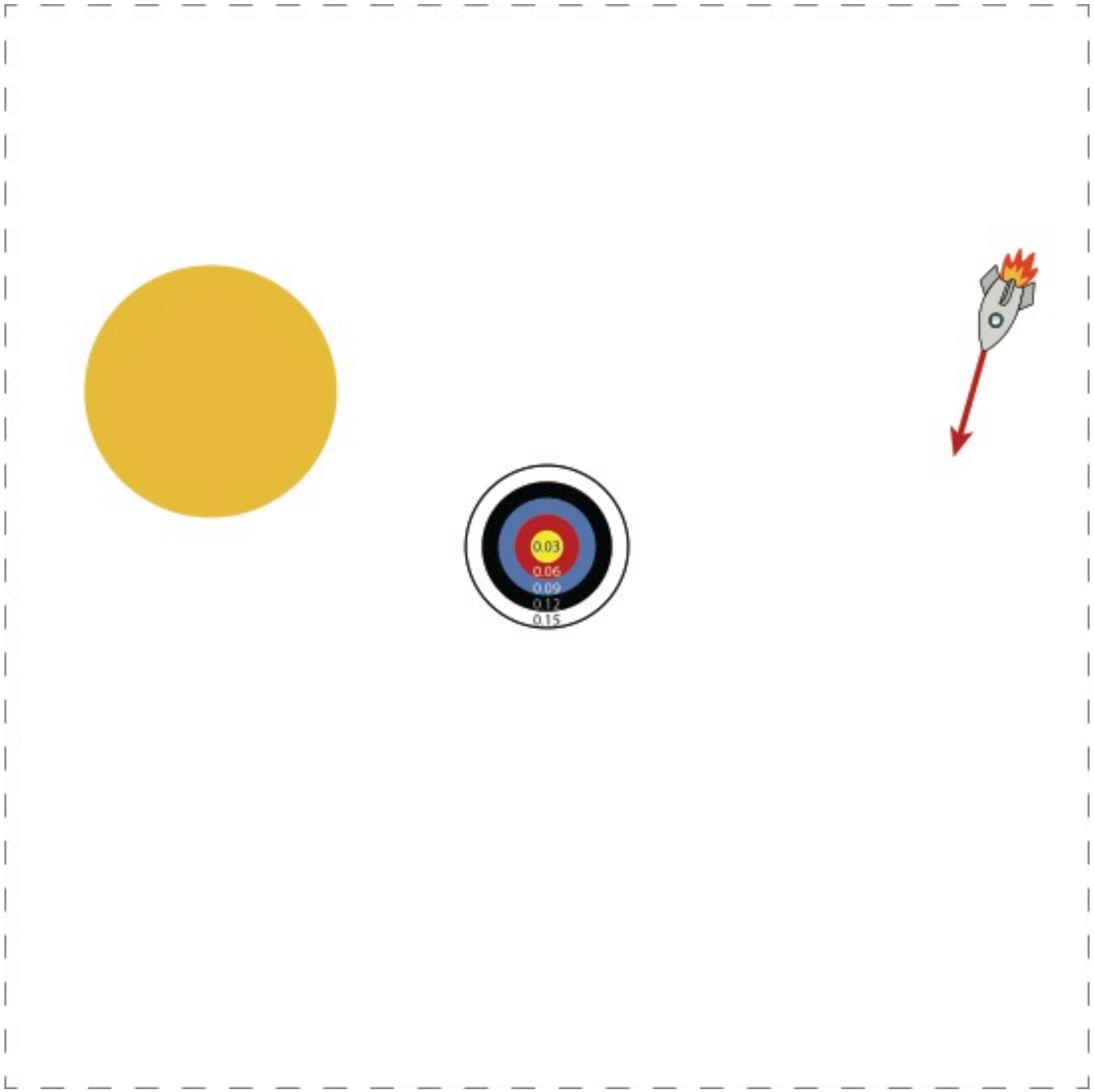
   Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018, arXiv)
   Bapst, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia, & Hamrick (2019, ICML)*

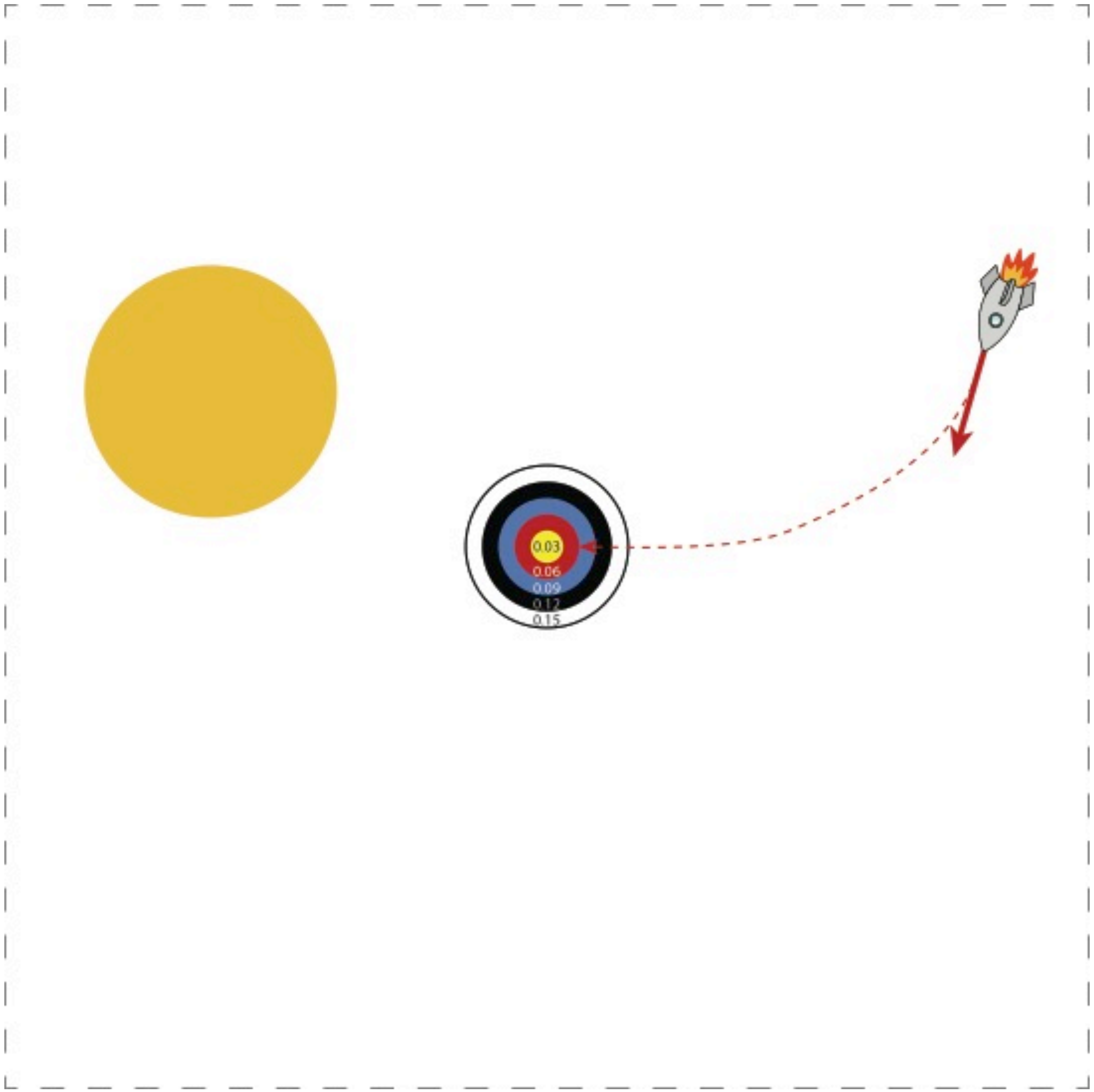


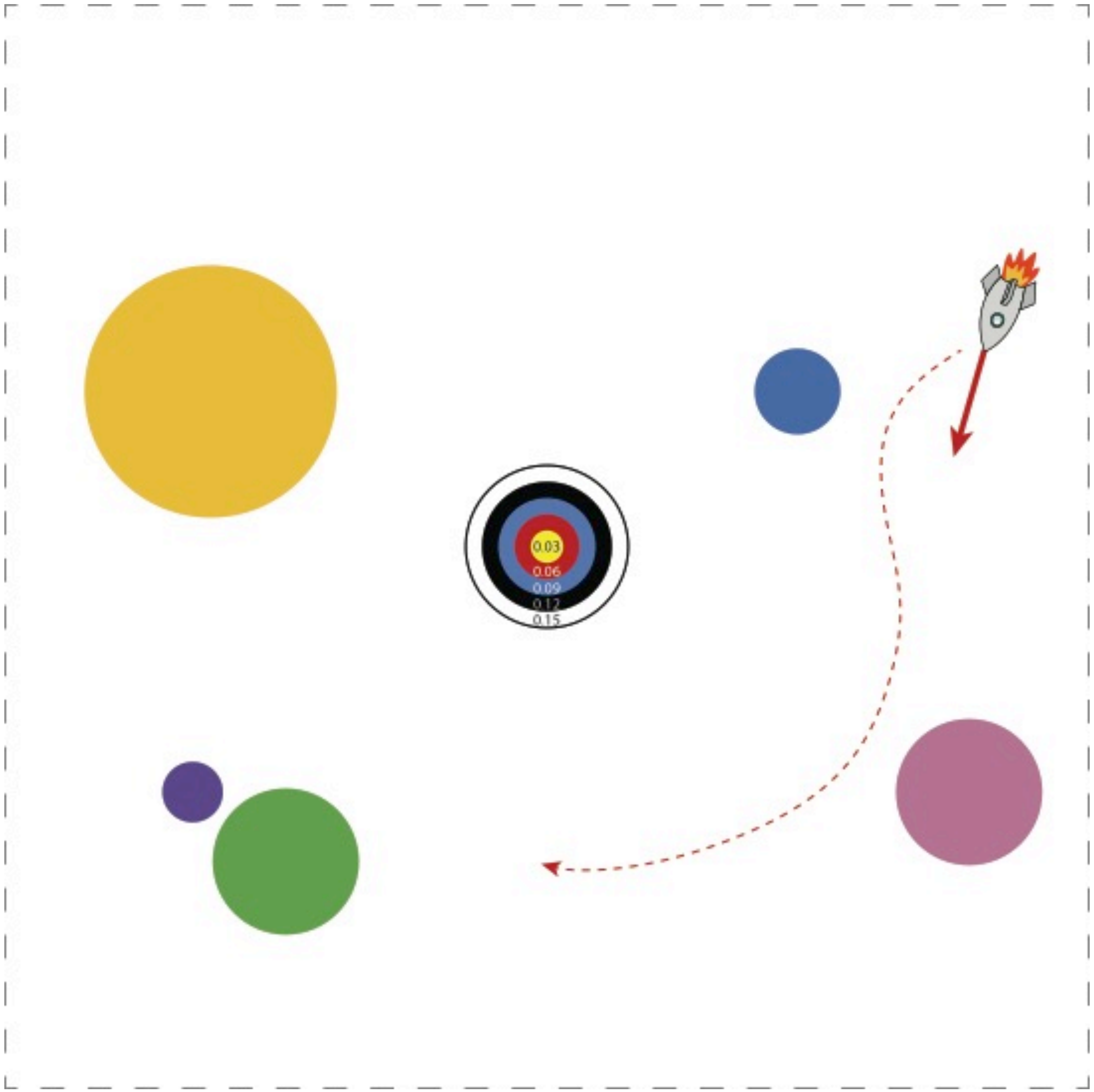




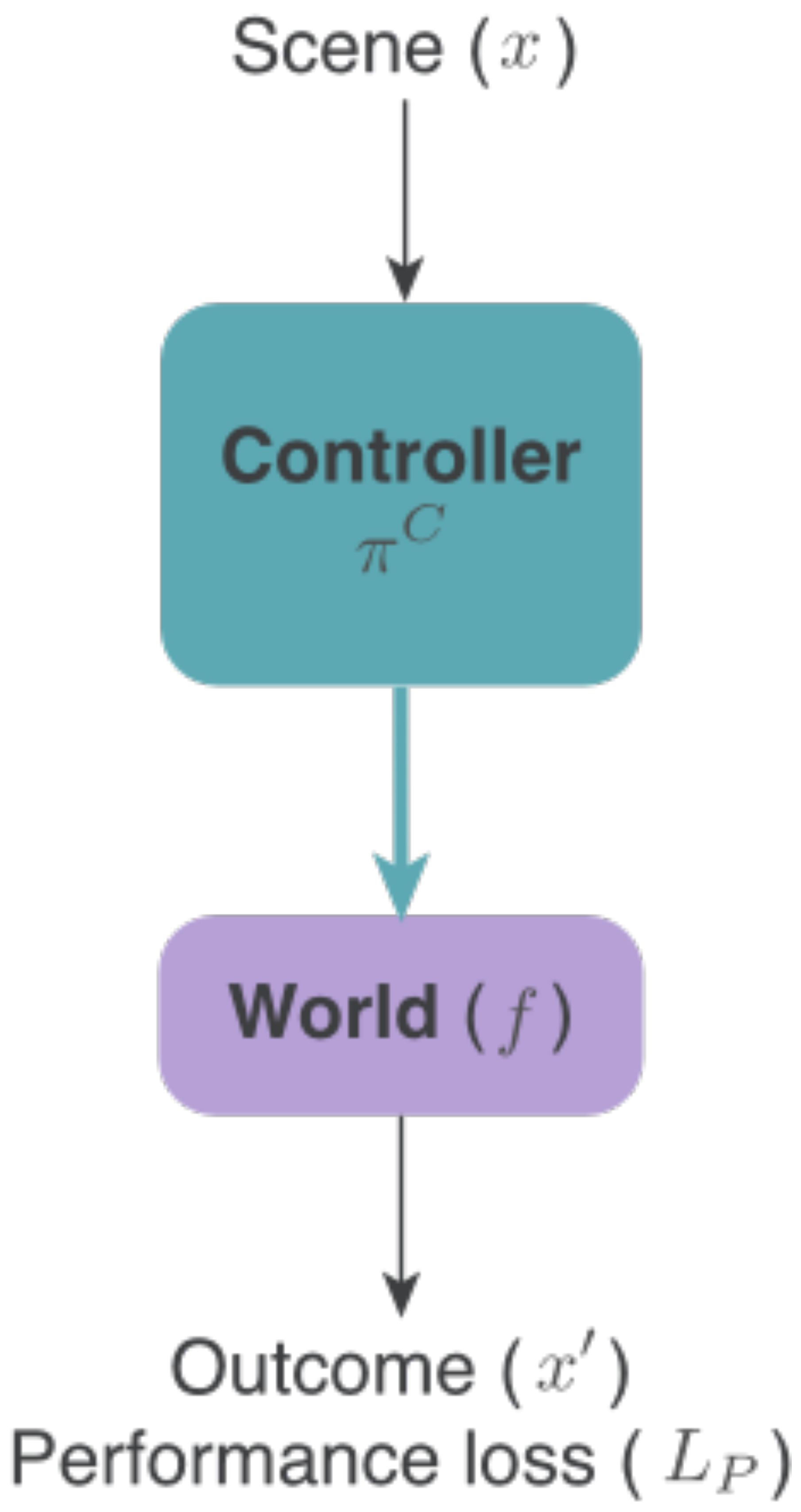
0.03
0.06
0.09
0.12
0.15


0.03
0.06
0.09
0.12
0.15


0.03
0.06
0.09
0.12
0.15

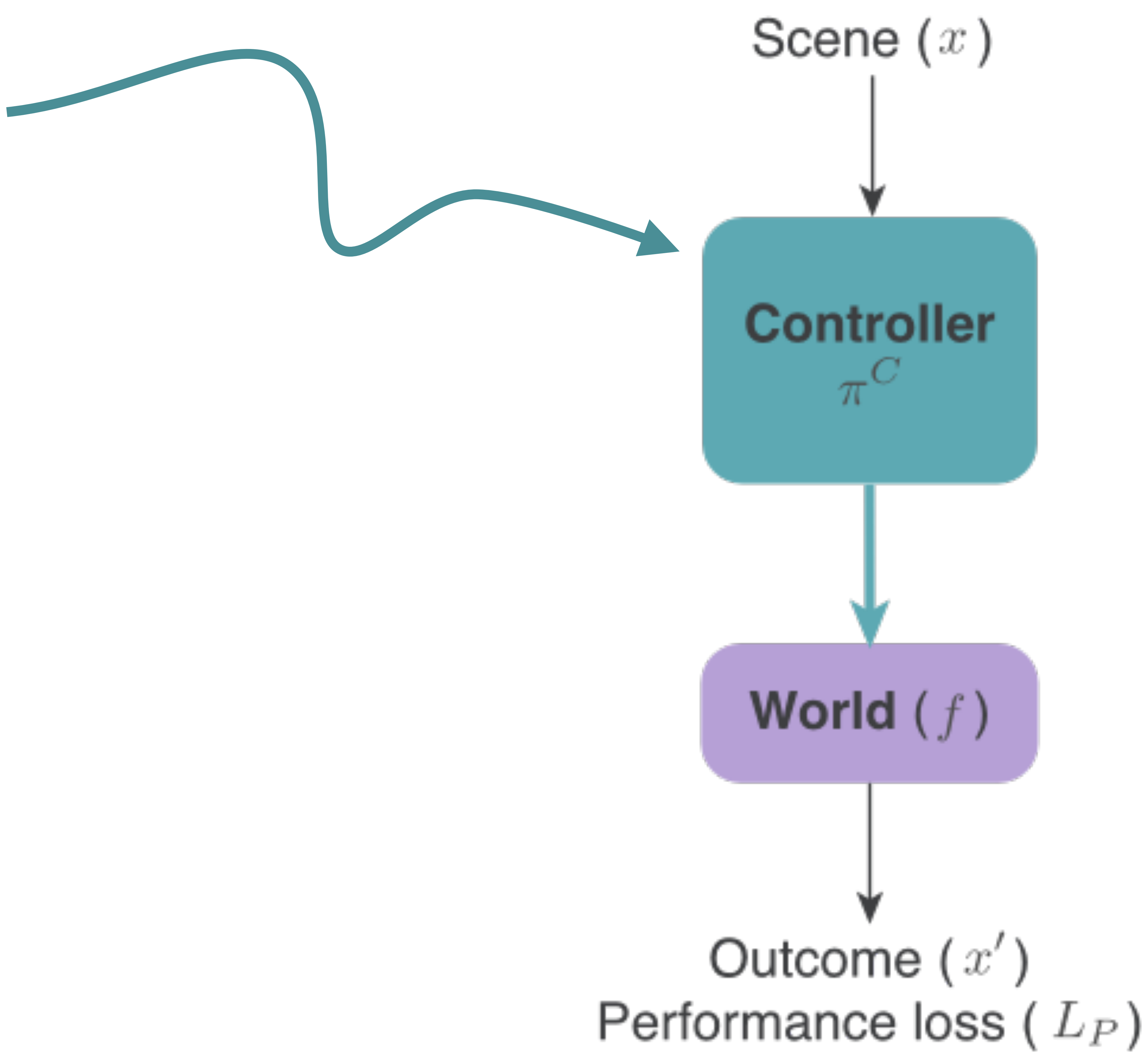

0.03
0.06
0.09
0.12
0.15

Reactive Controller



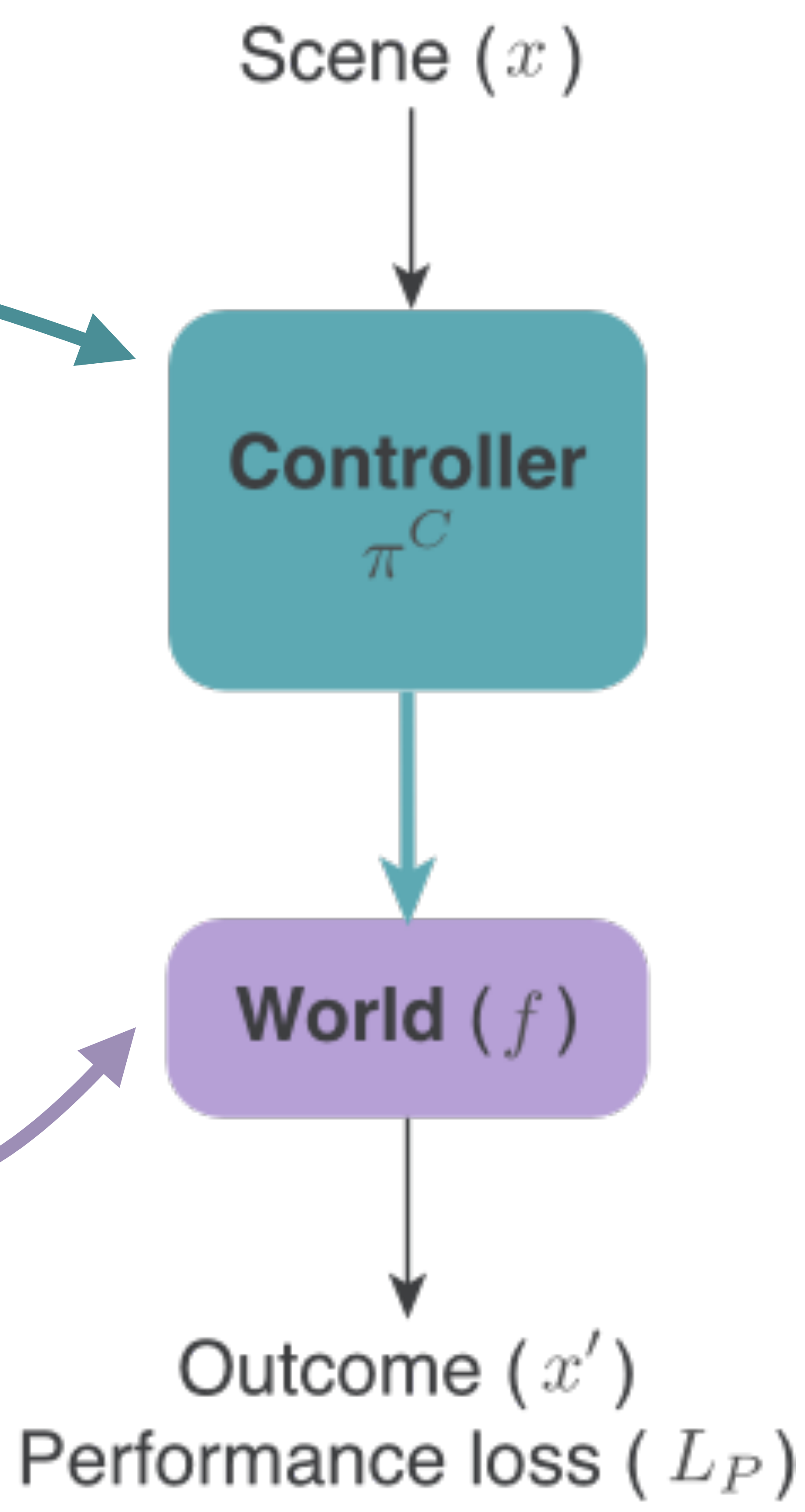
Reactive Controller

Controller (MLP)*:*
Proposes a control that is sent to the world to minimize performance loss



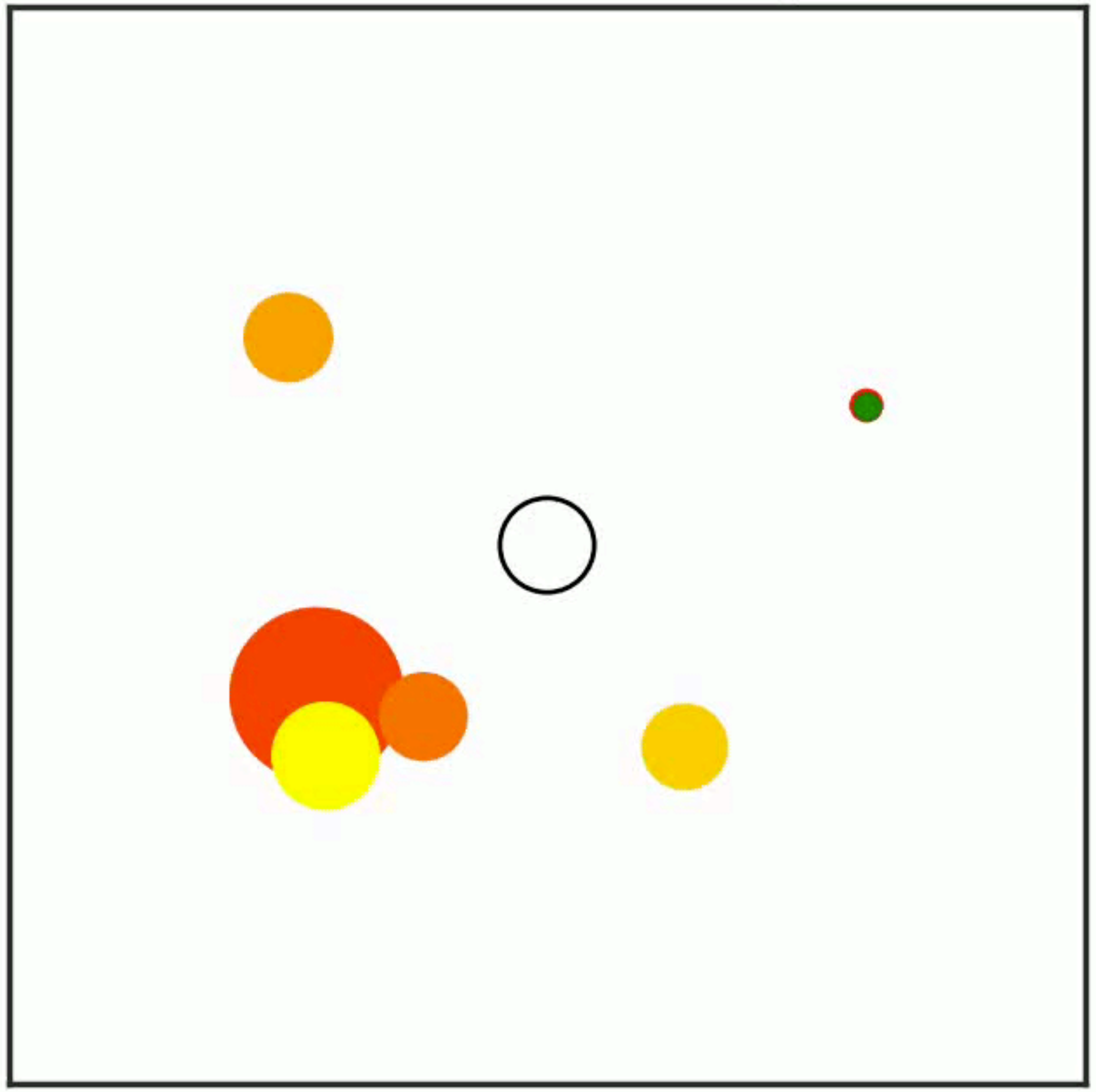
Reactive Controller

***Controller (MLP)**: Proposes a control that is sent to the world to minimize performance loss*

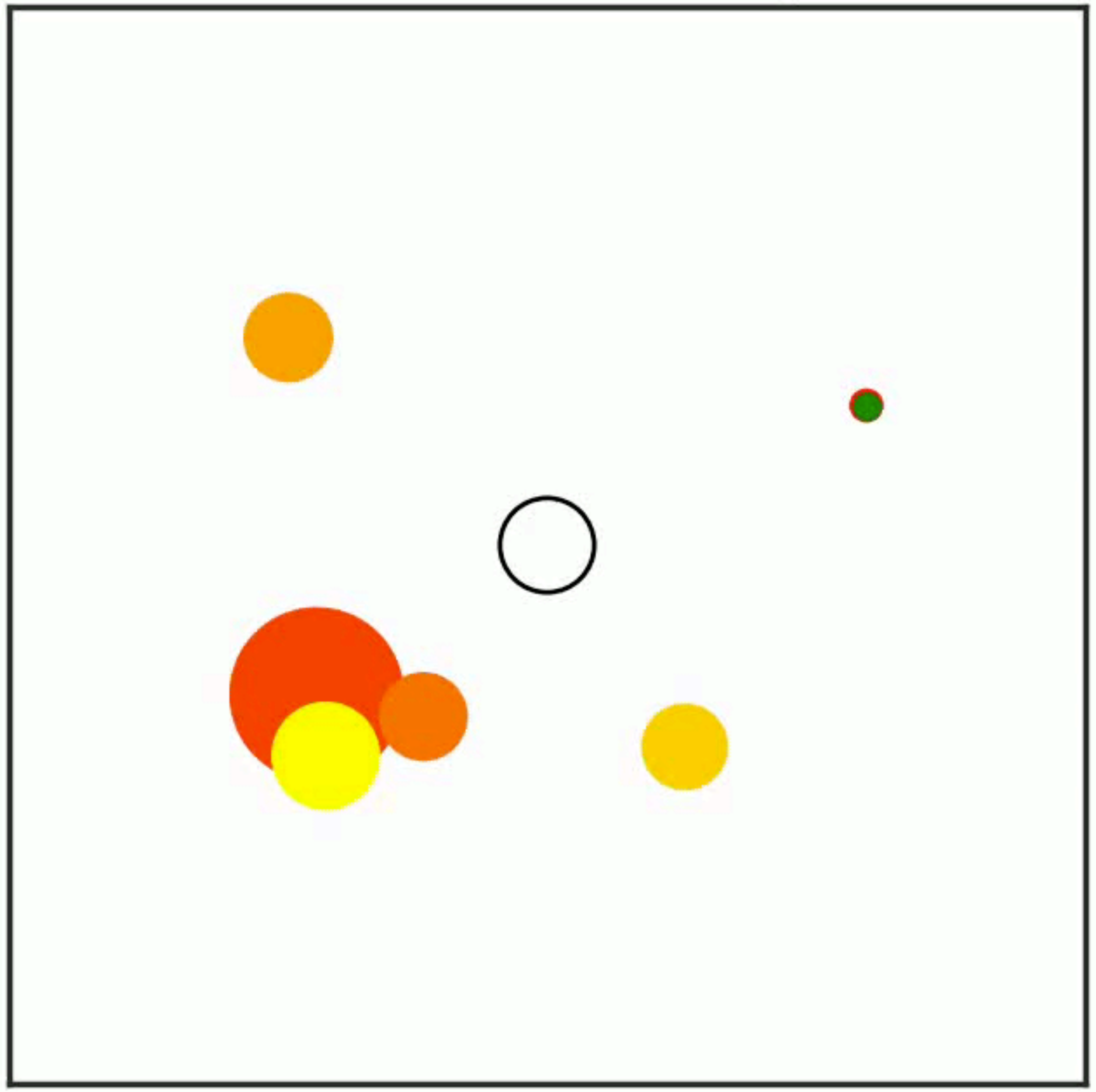


***World**: The true environment that the agent is acting in*

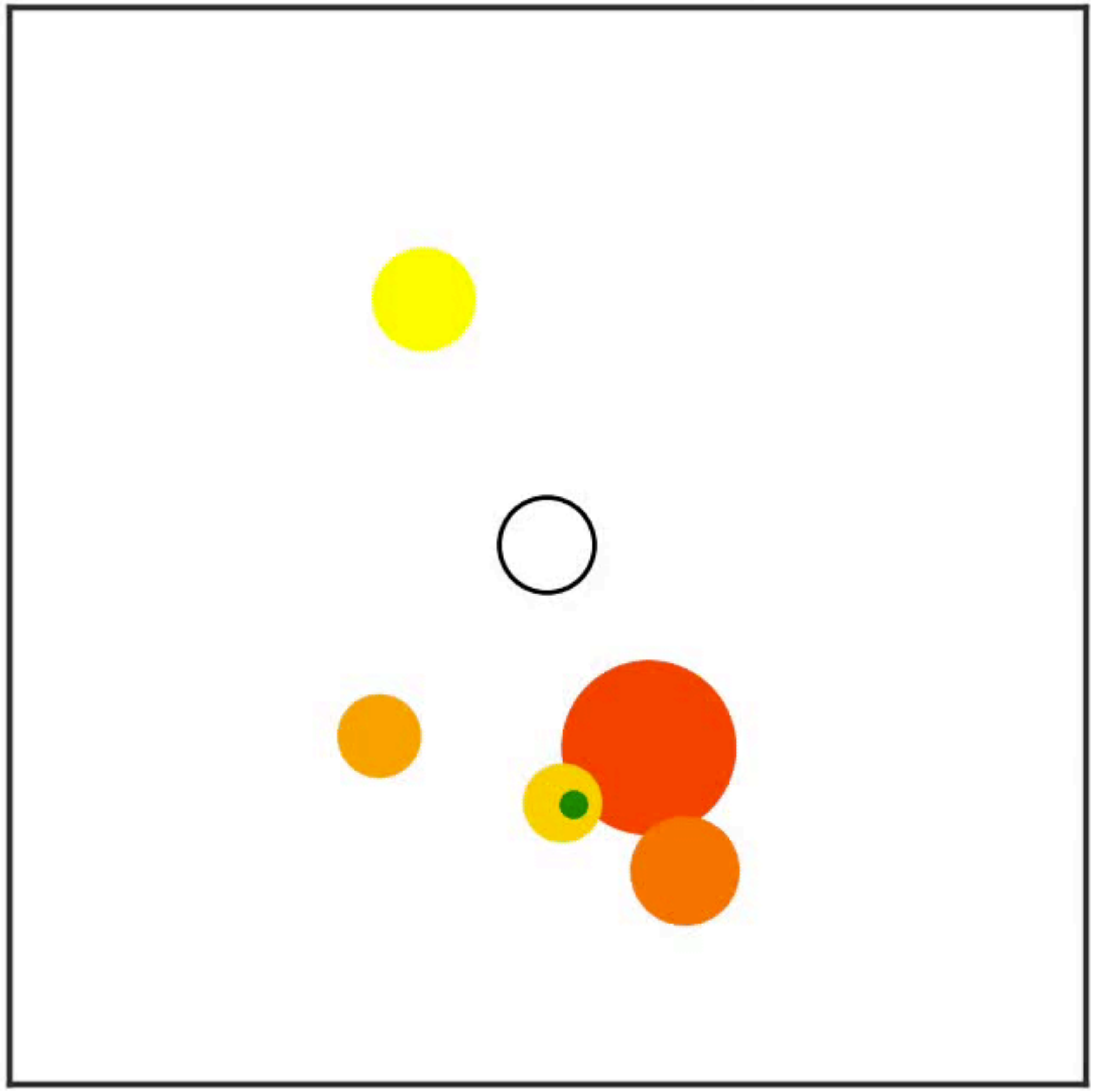
Reactive controller: simple task



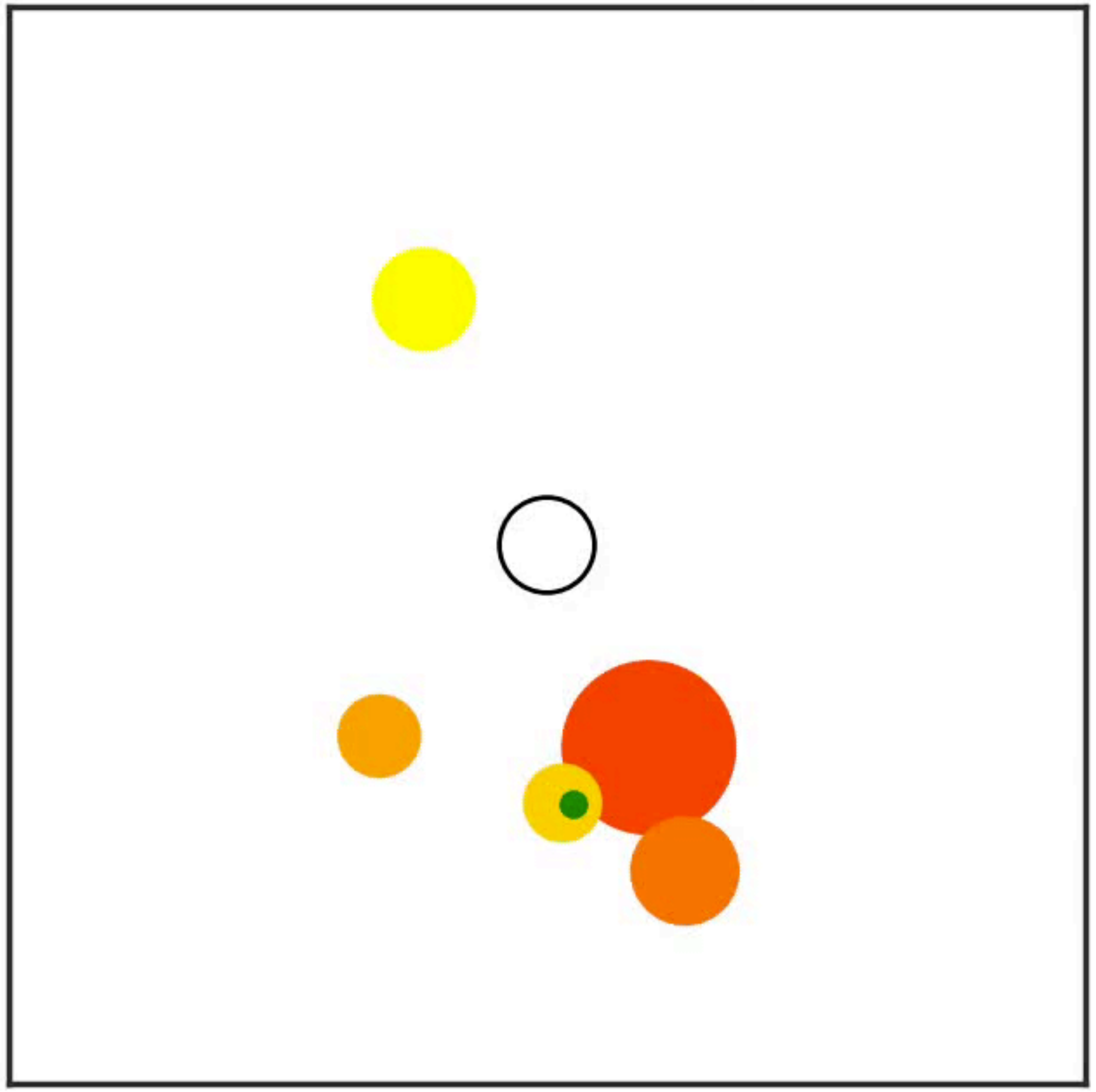
Reactive controller: simple task



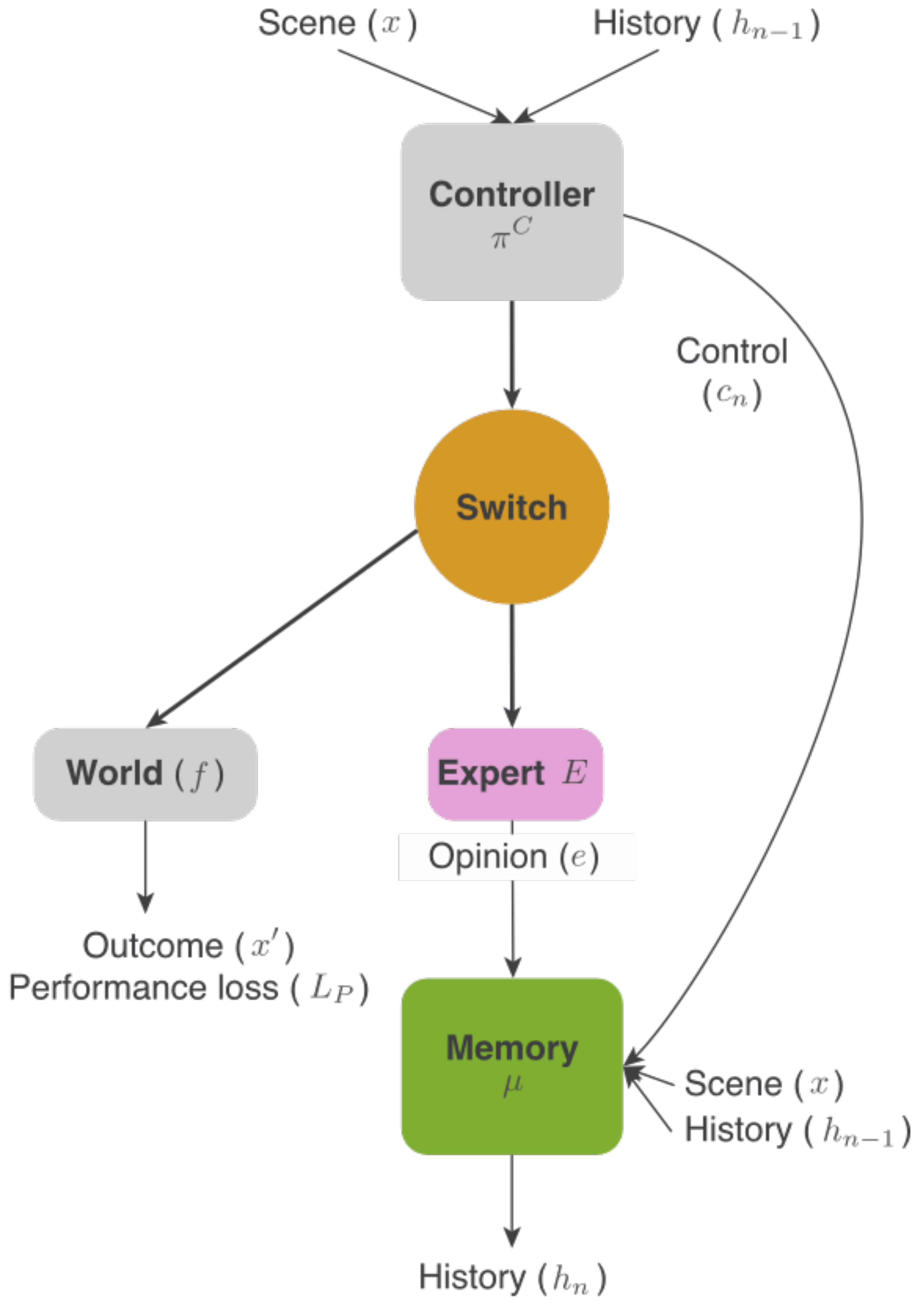
Reactive controller: hard task



Reactive controller: hard task

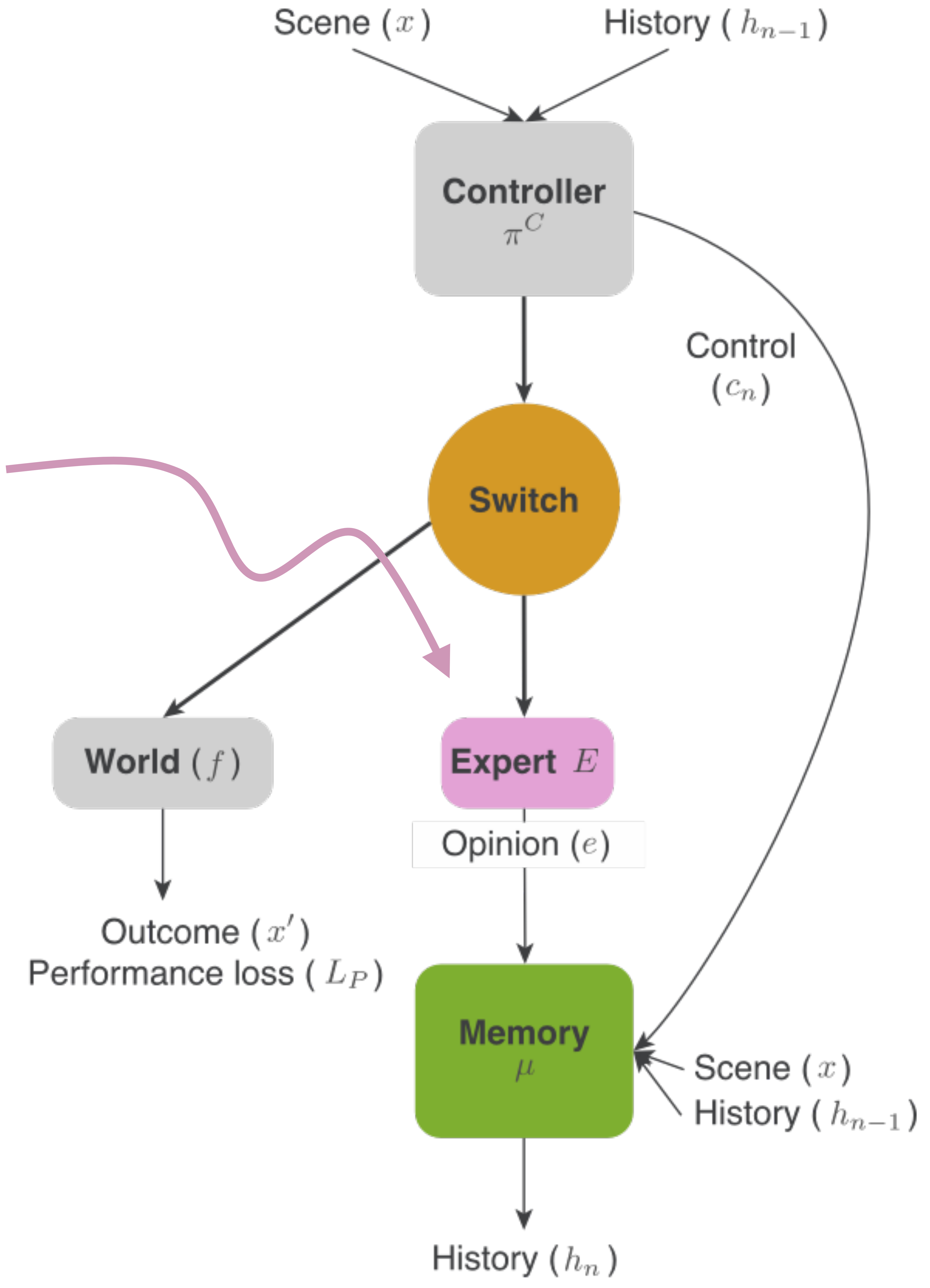


Iterative Controller



Iterative Controller

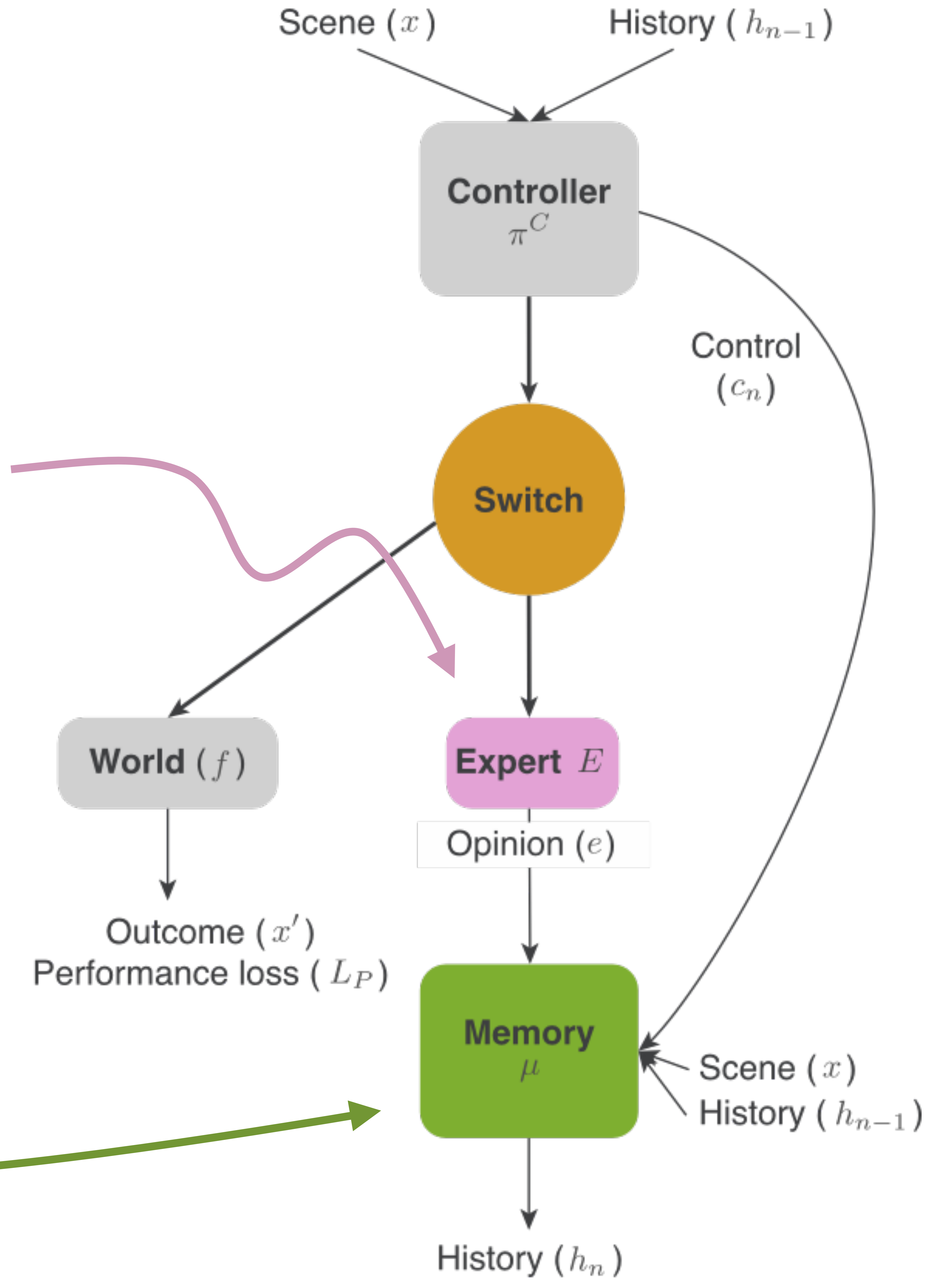
***Expert**:*
Model of the world that evaluates proposed controls



Iterative Controller

***Expert**:*
Model of the world that evaluates proposed controls

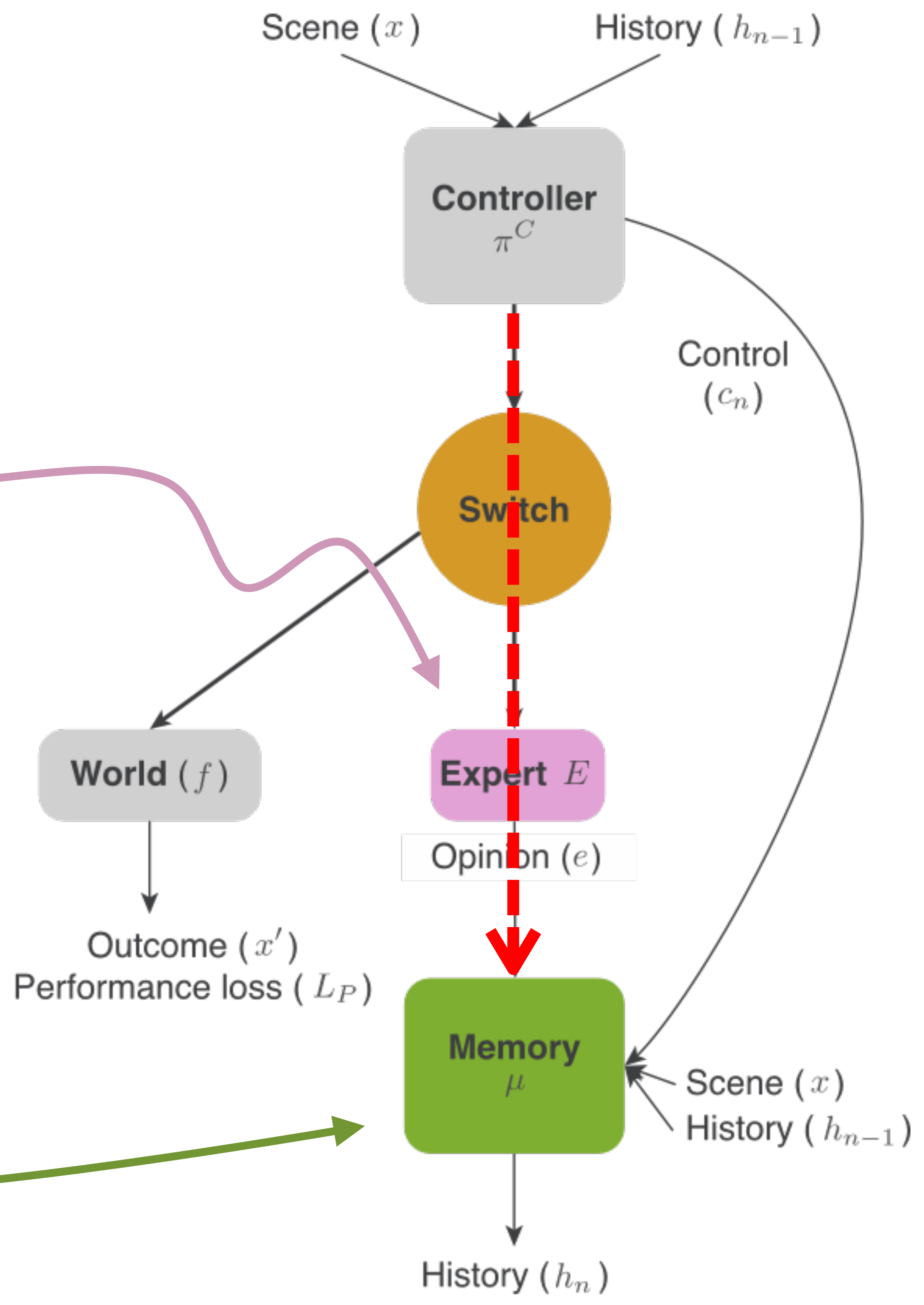
***Memory (LSTM)**:*
Encodes the full history of controls and opinions



Iterative Controller

***Expert**:*
Model of the world that evaluates proposed controls

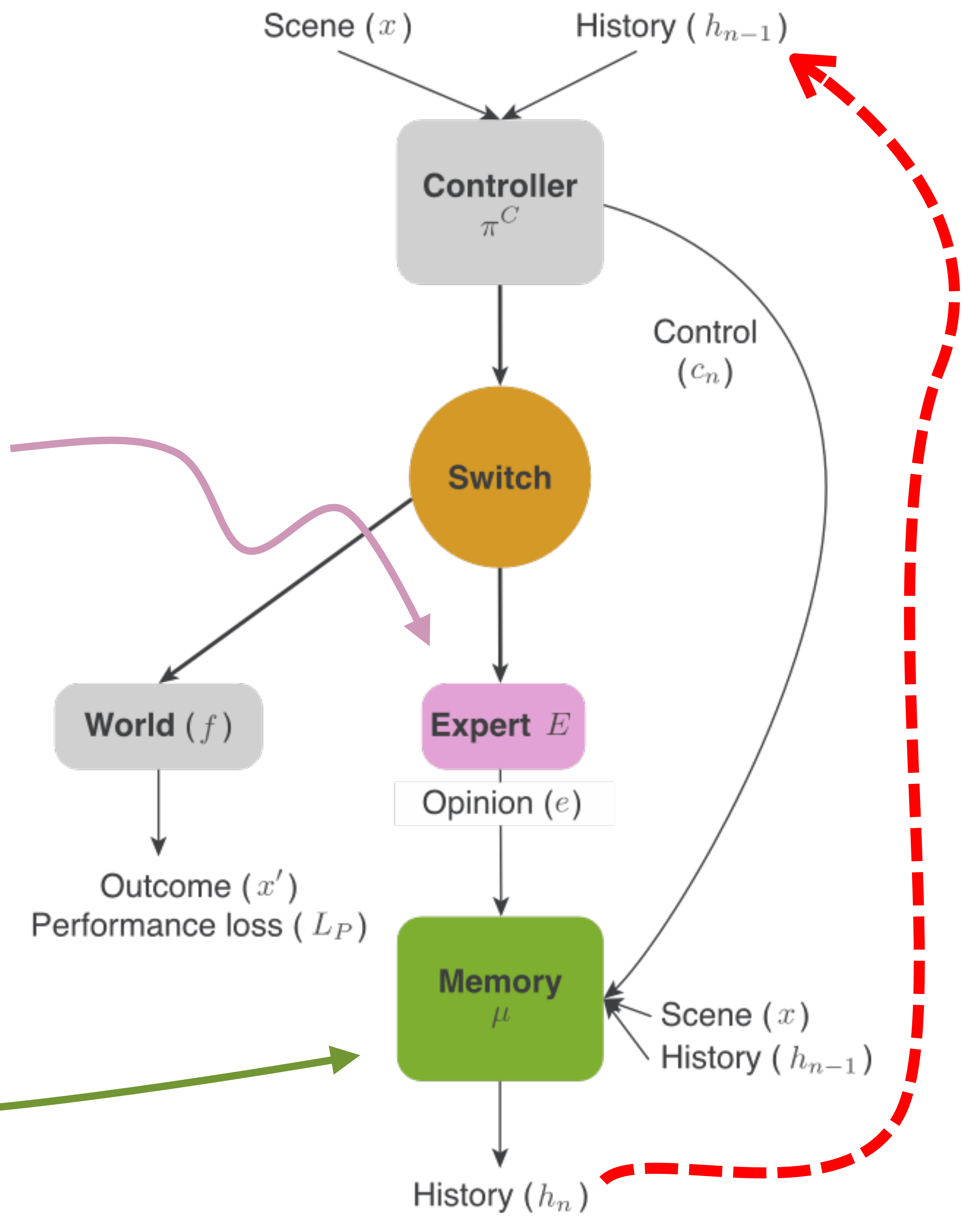
***Memory (LSTM)**:*
Encodes the full history of controls and opinions



Iterative Controller

***Expert**:*
Model of the world that evaluates proposed controls

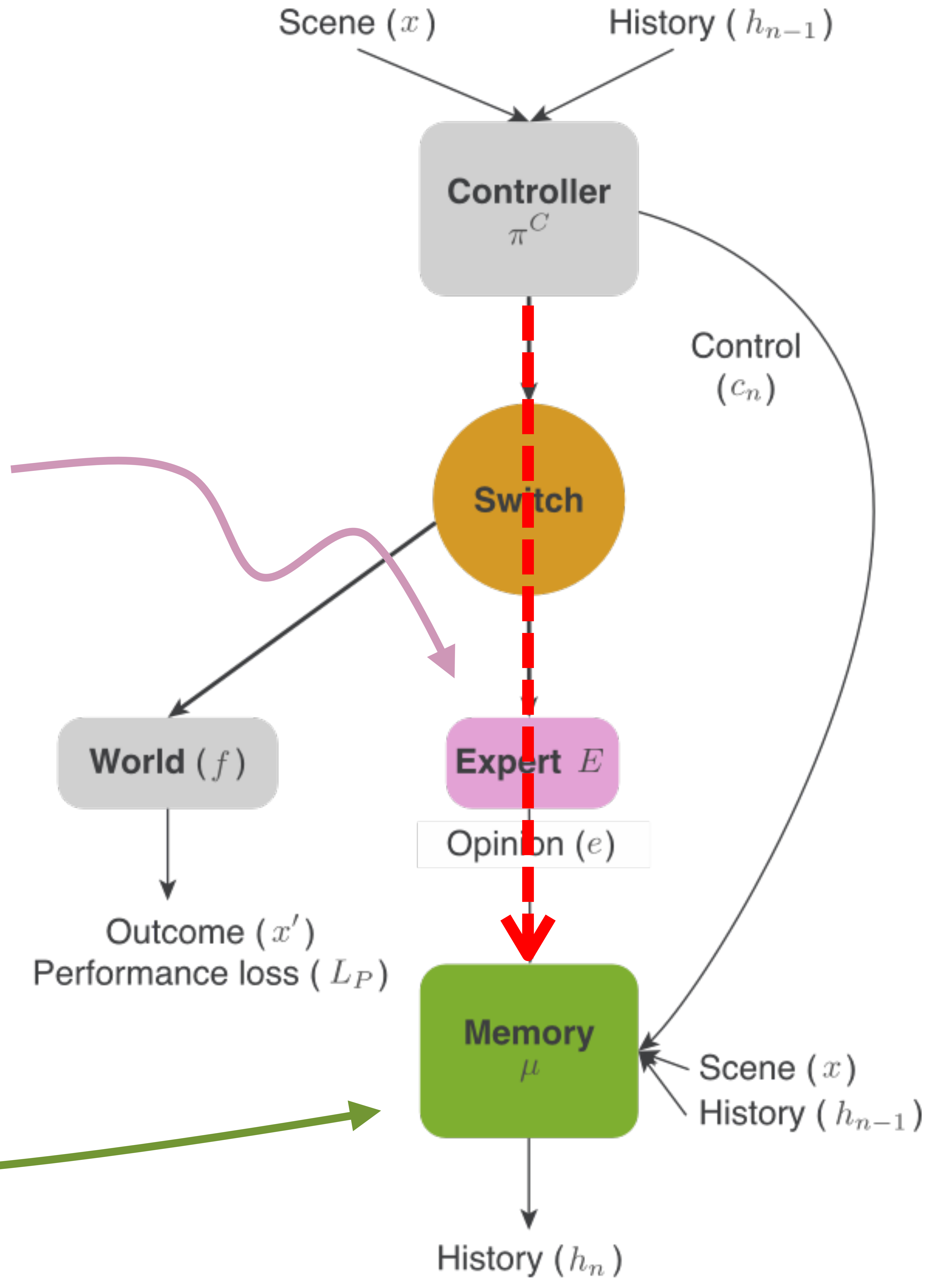
***Memory (LSTM)**:*
Encodes the full history of controls and opinions



Iterative Controller

Expert:
Model of the world that evaluates proposed controls

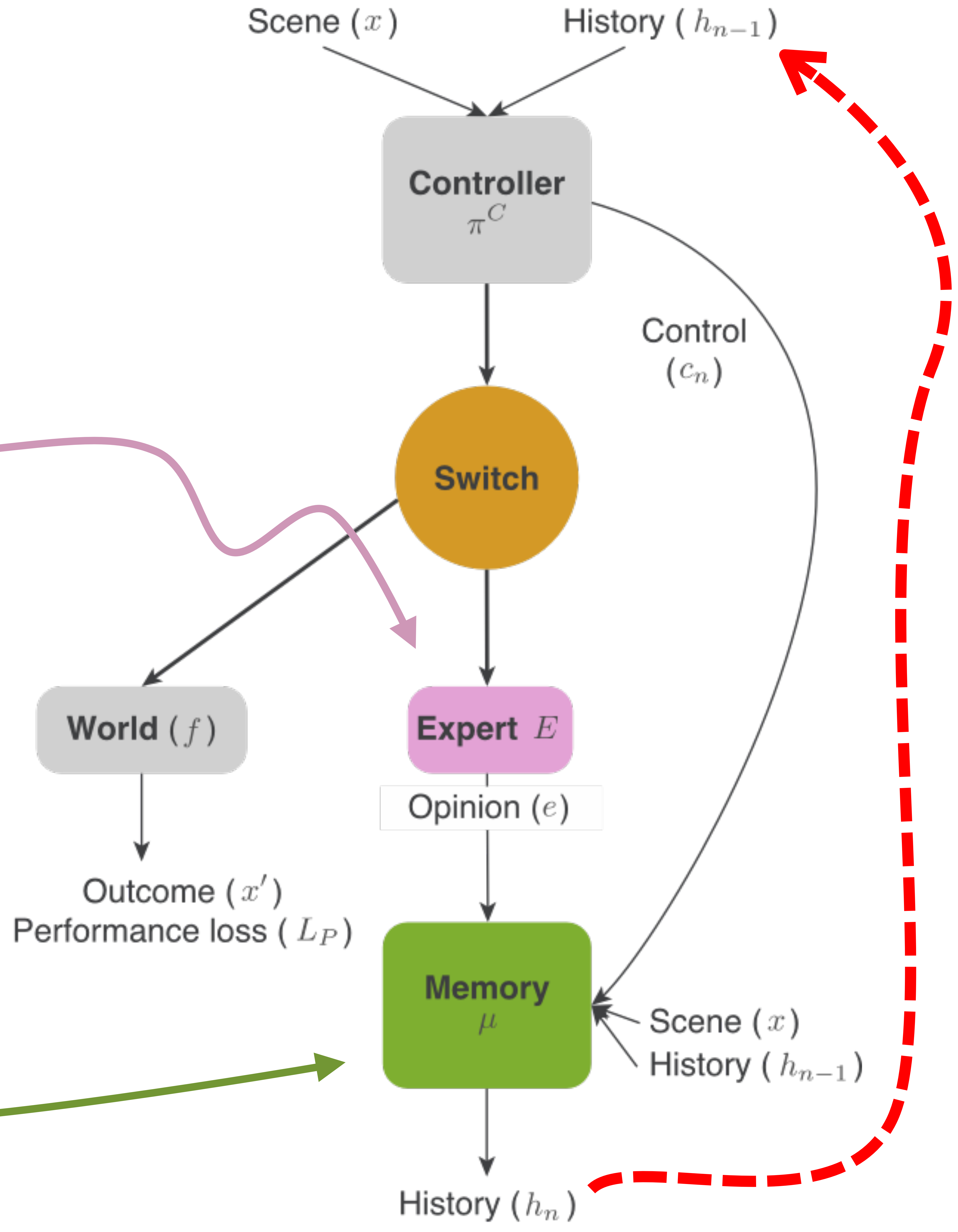
Memory (LSTM):
Encodes the full history of controls and opinions



Iterative Controller

Expert:
Model of the world that evaluates proposed controls

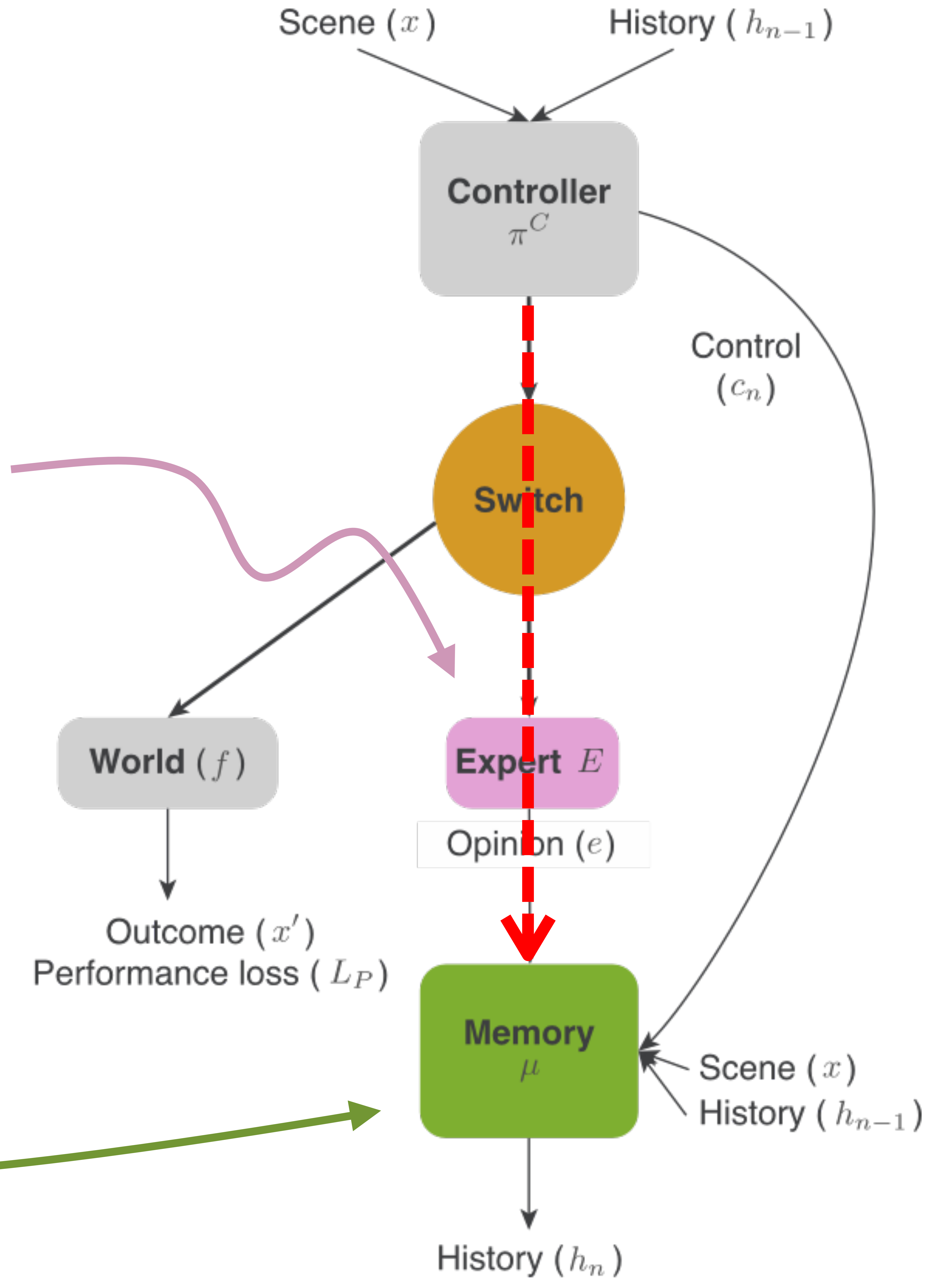
Memory (LSTM):
Encodes the full history of controls and opinions



Iterative Controller

***Expert**:*
Model of the world that evaluates proposed controls

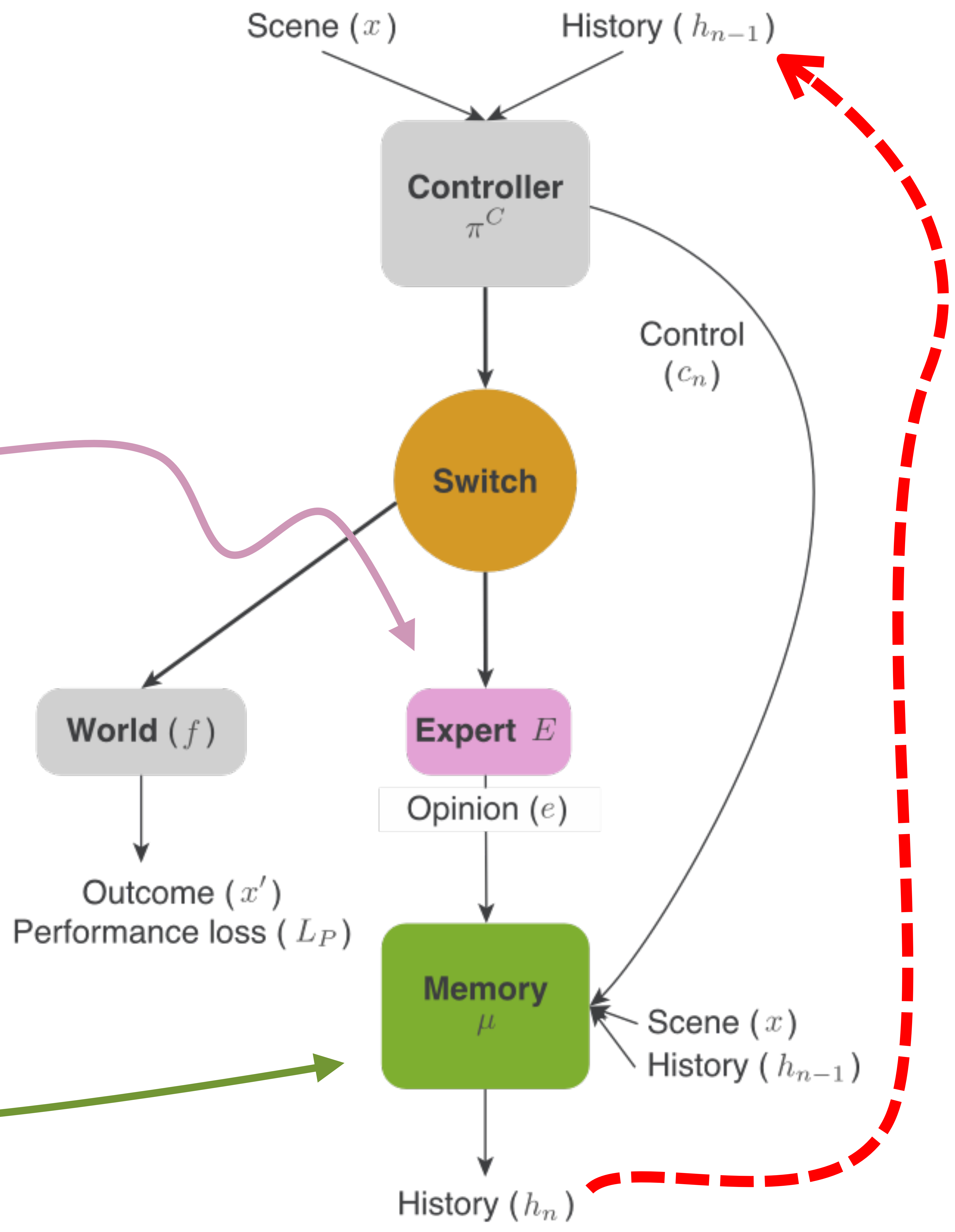
***Memory (LSTM)**:*
Encodes the full history of controls and opinions



Iterative Controller

Expert: *Model of the world that evaluates proposed controls*

Memory (LSTM): *Encodes the full history of controls and opinions*



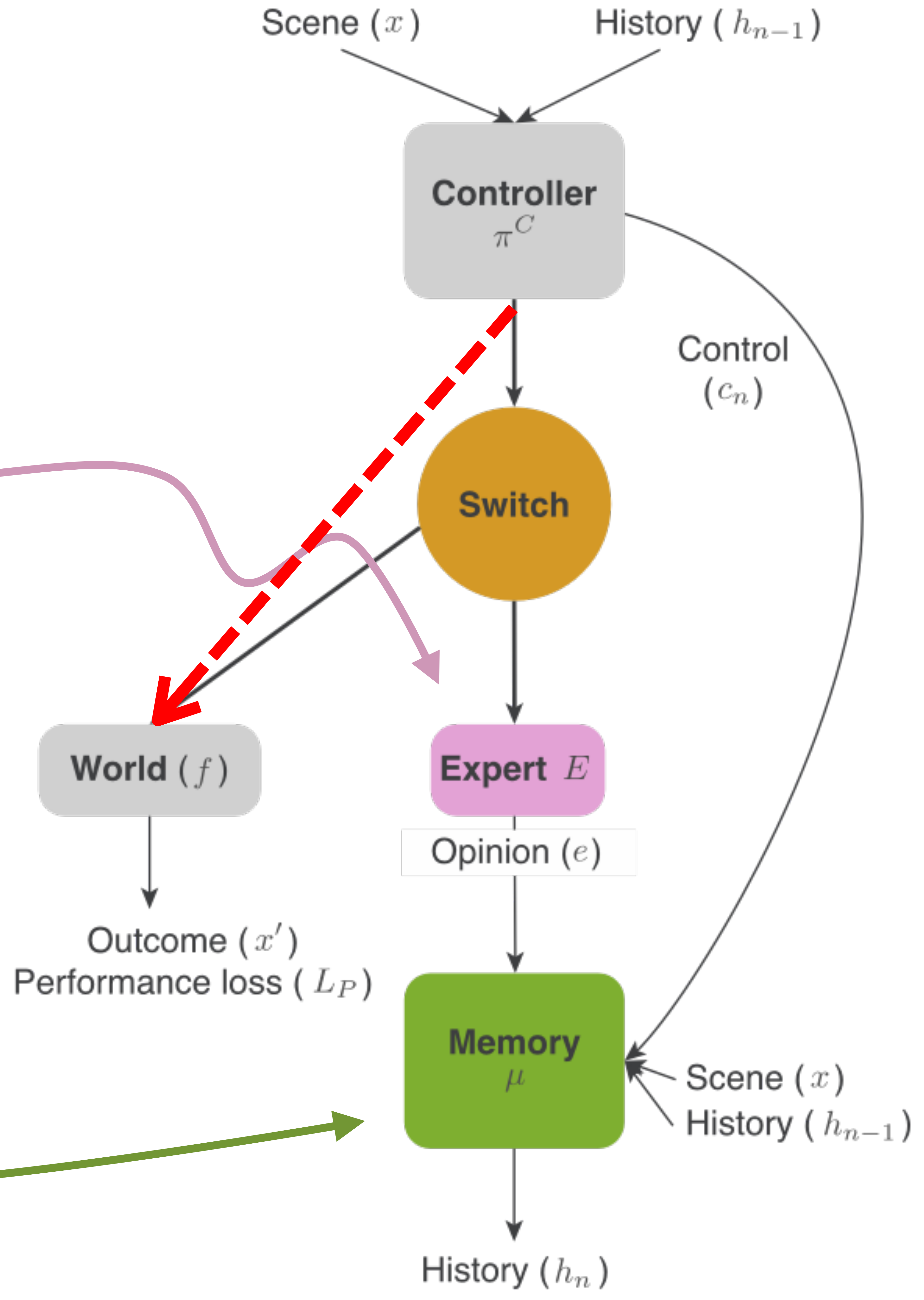
DeepMind

Jessica Hamrick (@jhamrick)

Iterative Controller

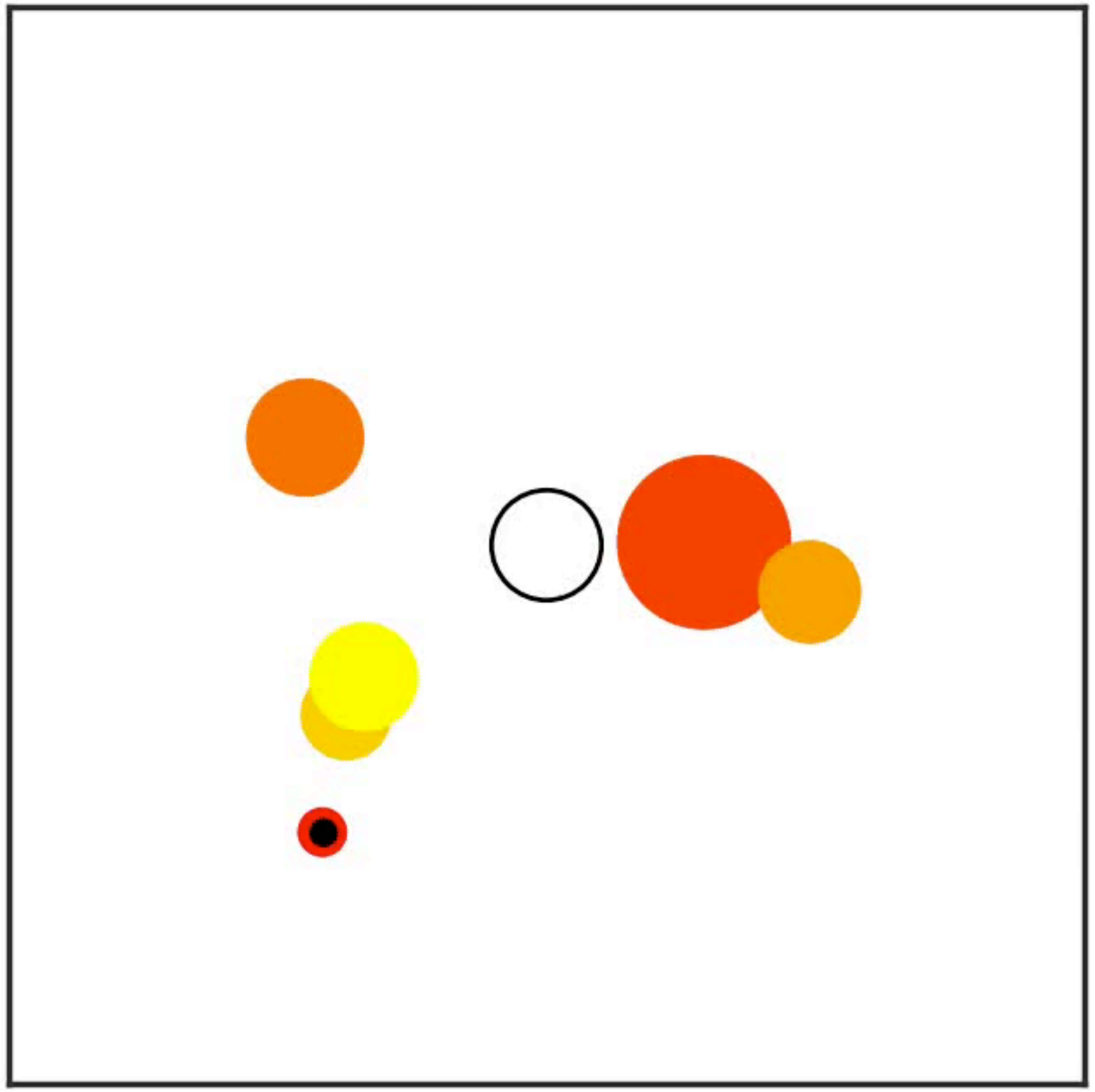
***Expert**:*
Model of the world that evaluates proposed controls

***Memory (LSTM)**:*
Encodes the full history of controls and opinions

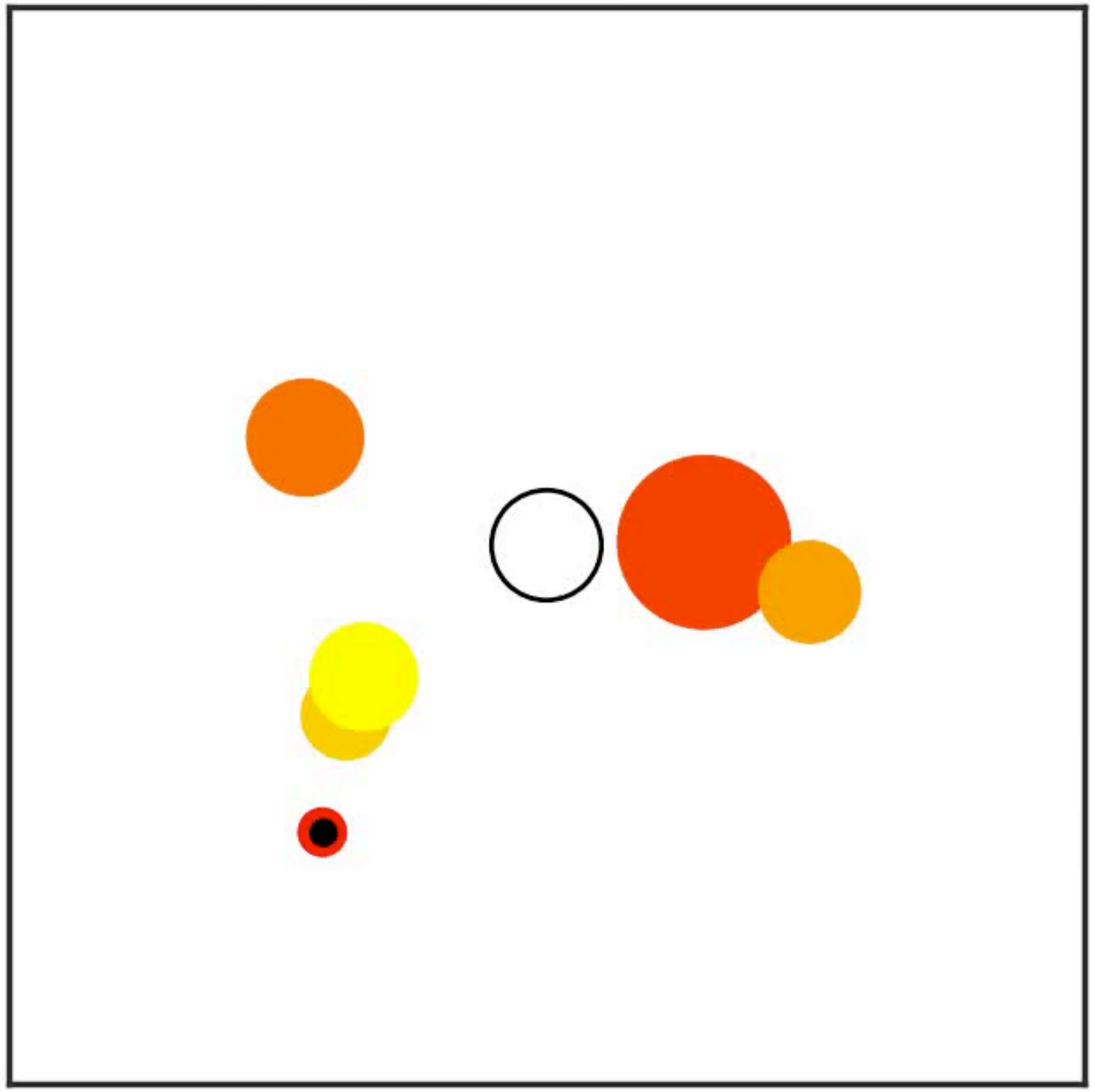


DeepMind

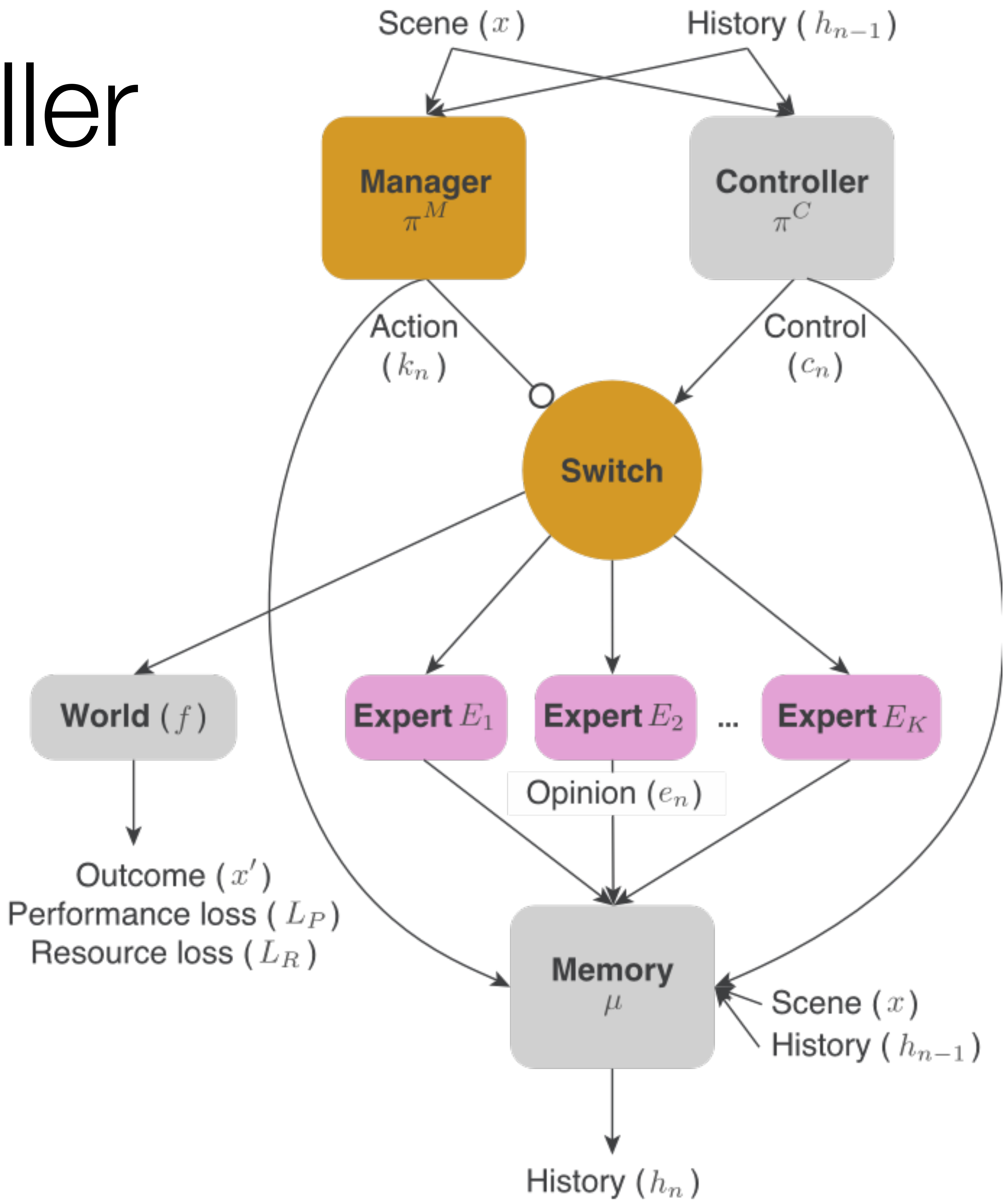
Iterative controller



Iterative controller

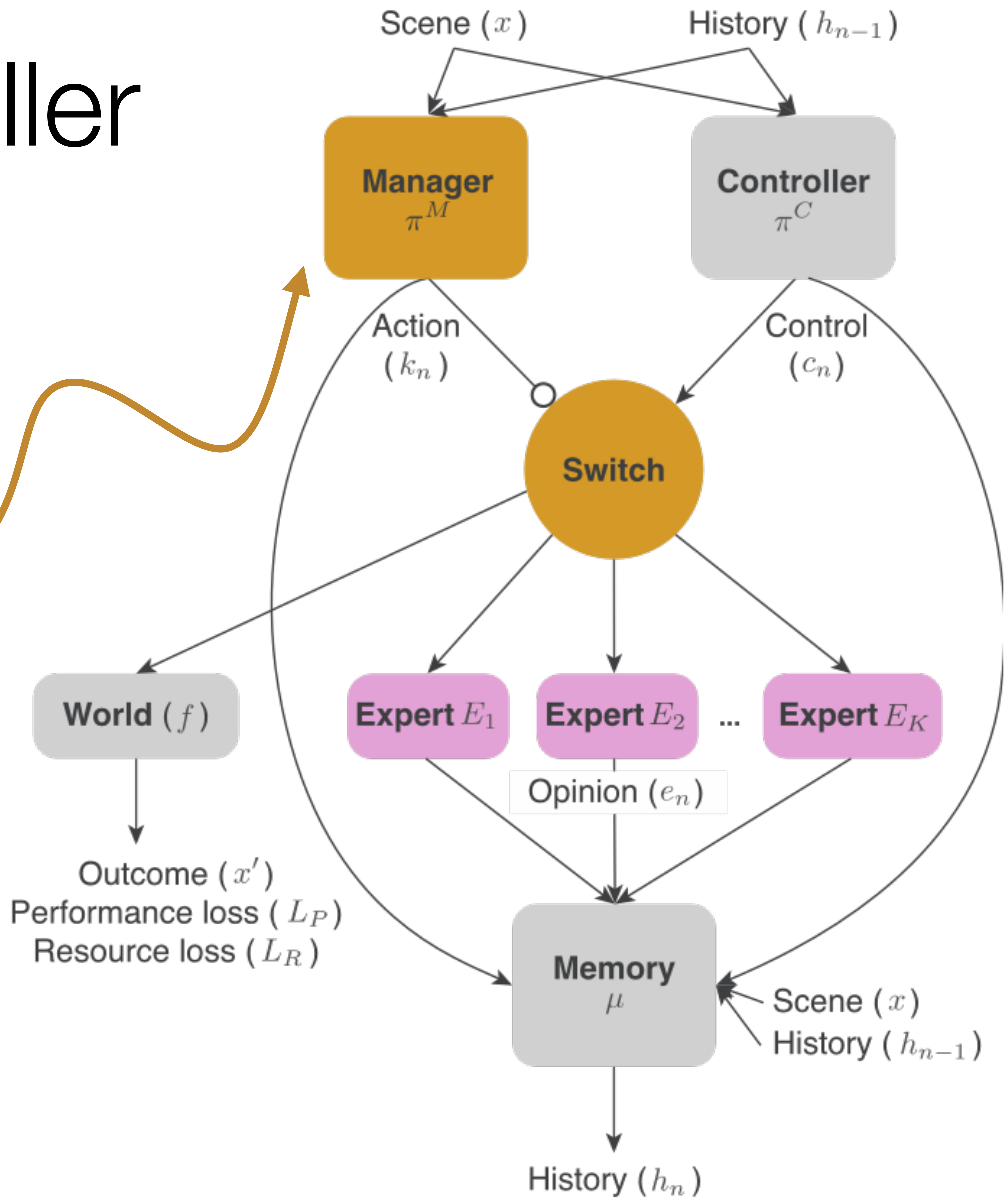


Metacontroller



Metacontroller

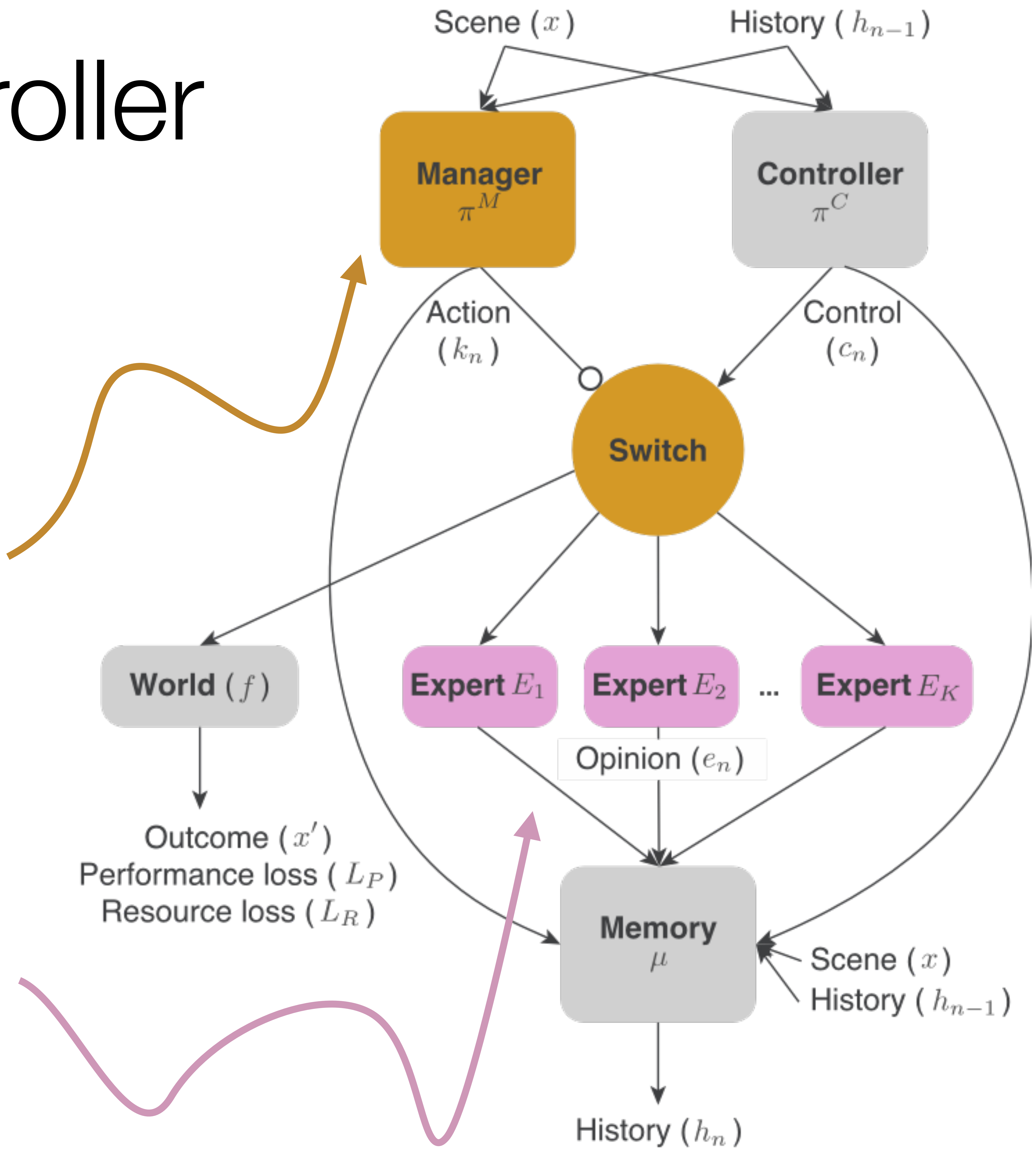
Manager (MLP)**: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to **minimize performance loss + resource loss



Metacontroller

Manager (MLP)**: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to **minimize performance loss + resource loss

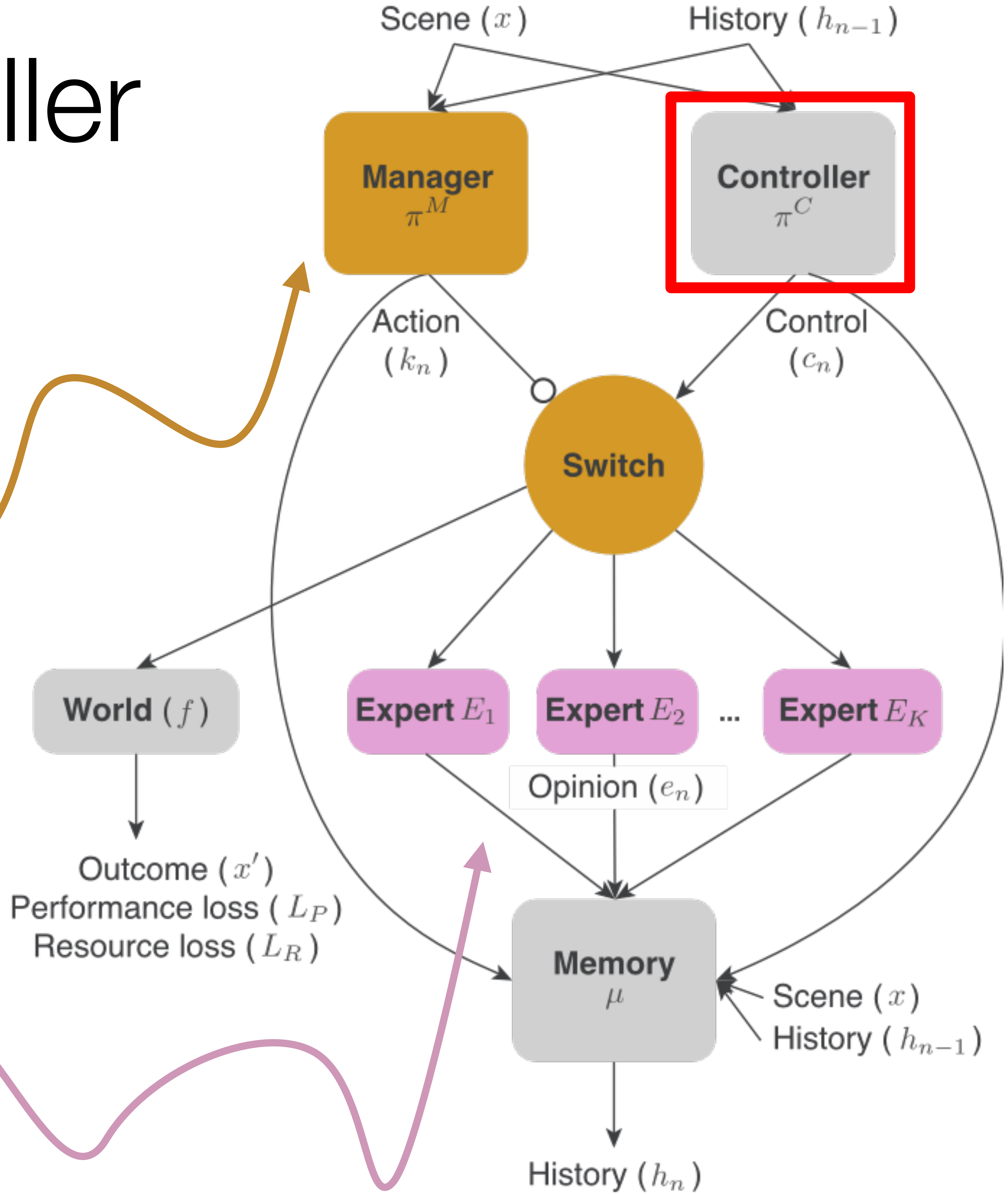
Experts**: Different models of the world, each with **different resource costs



Metacontroller

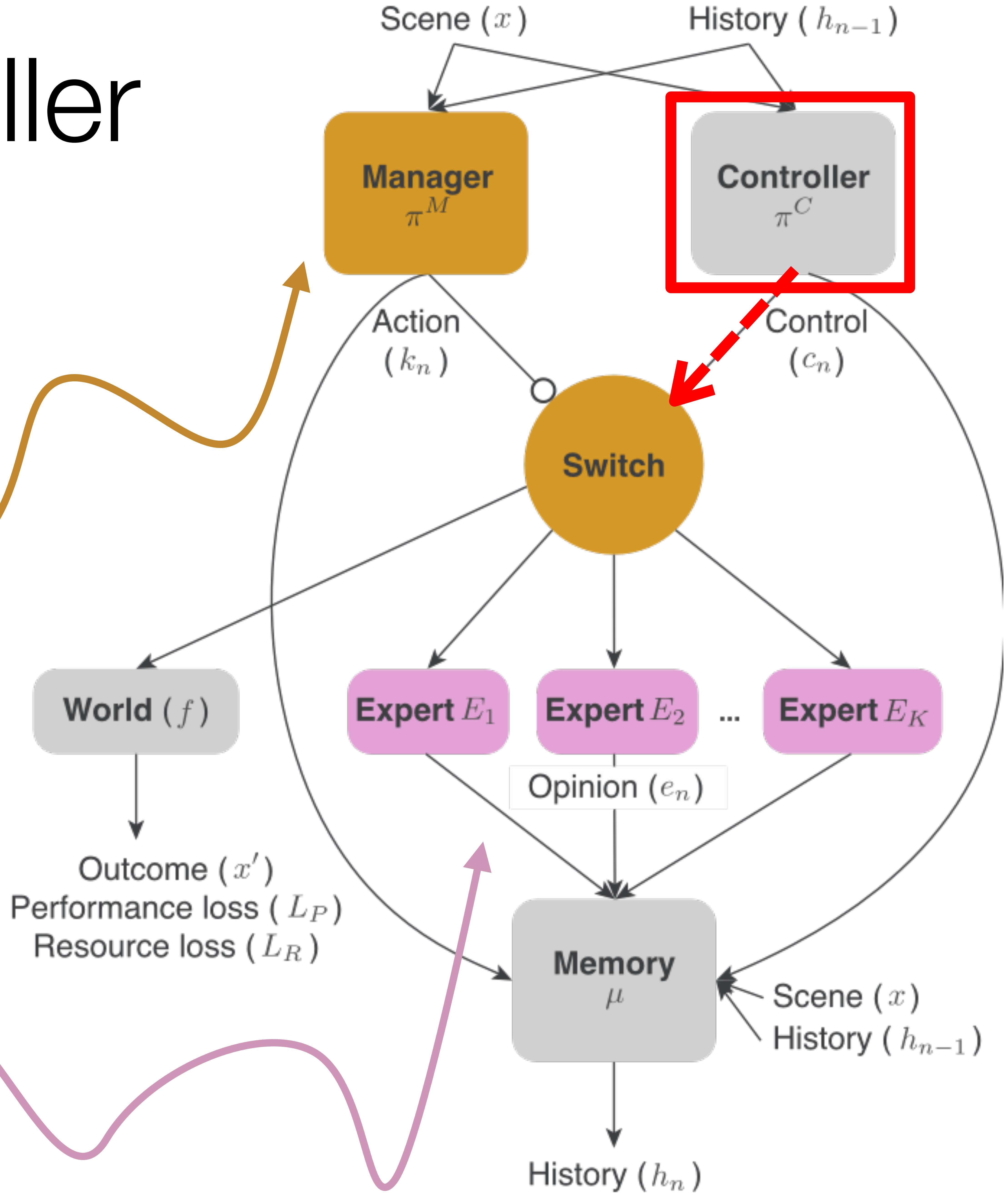
Manager (MLP)**: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to **minimize performance loss + resource loss

Experts**: Different models of the world, each with **different resource costs



Metacontroller

Manager (MLP)**: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to **minimize performance loss + resource loss

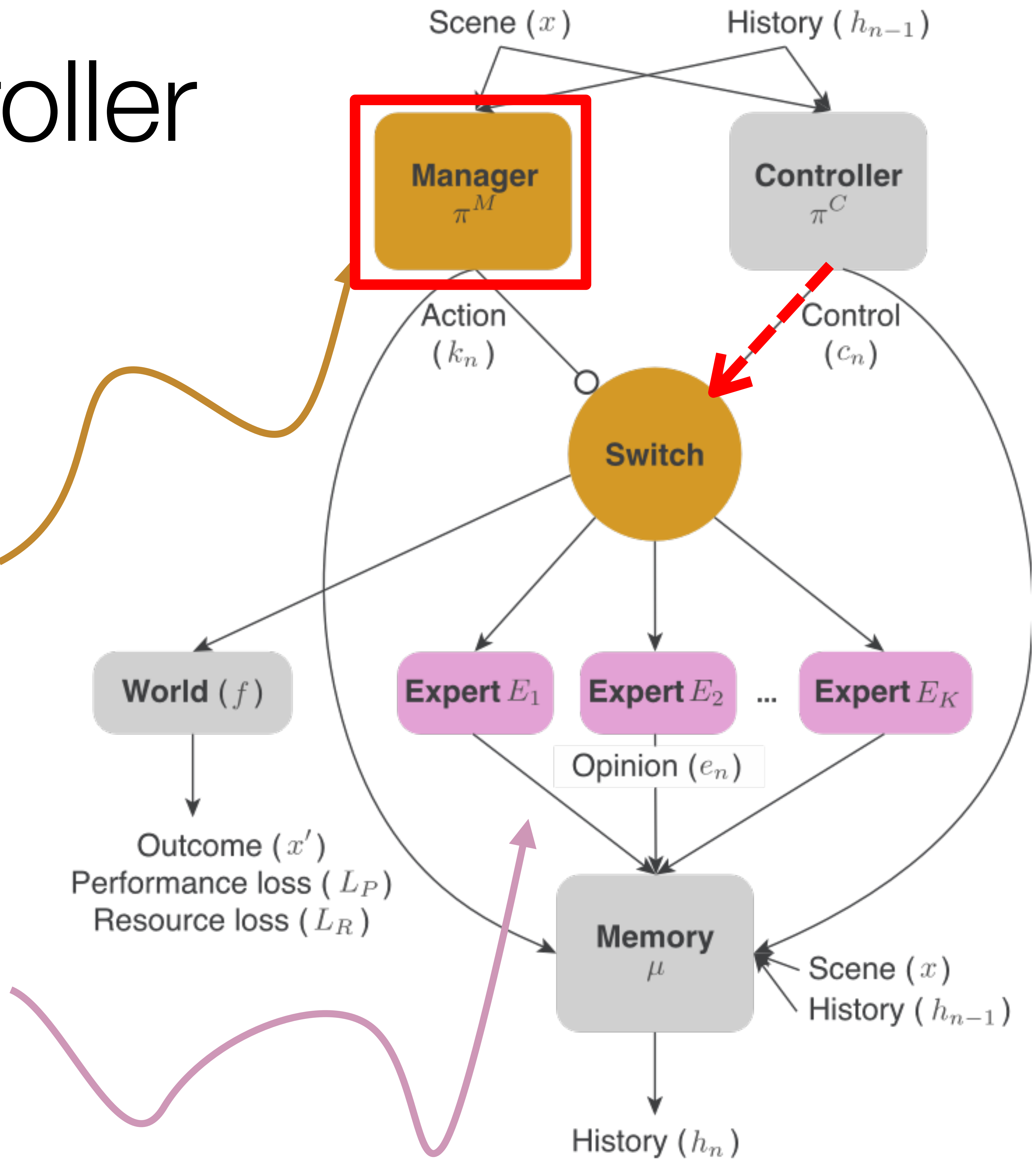


Experts**: Different models of the world, each with **different resource costs

Metacontroller

Manager (MLP)**: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to **minimize performance loss + resource loss

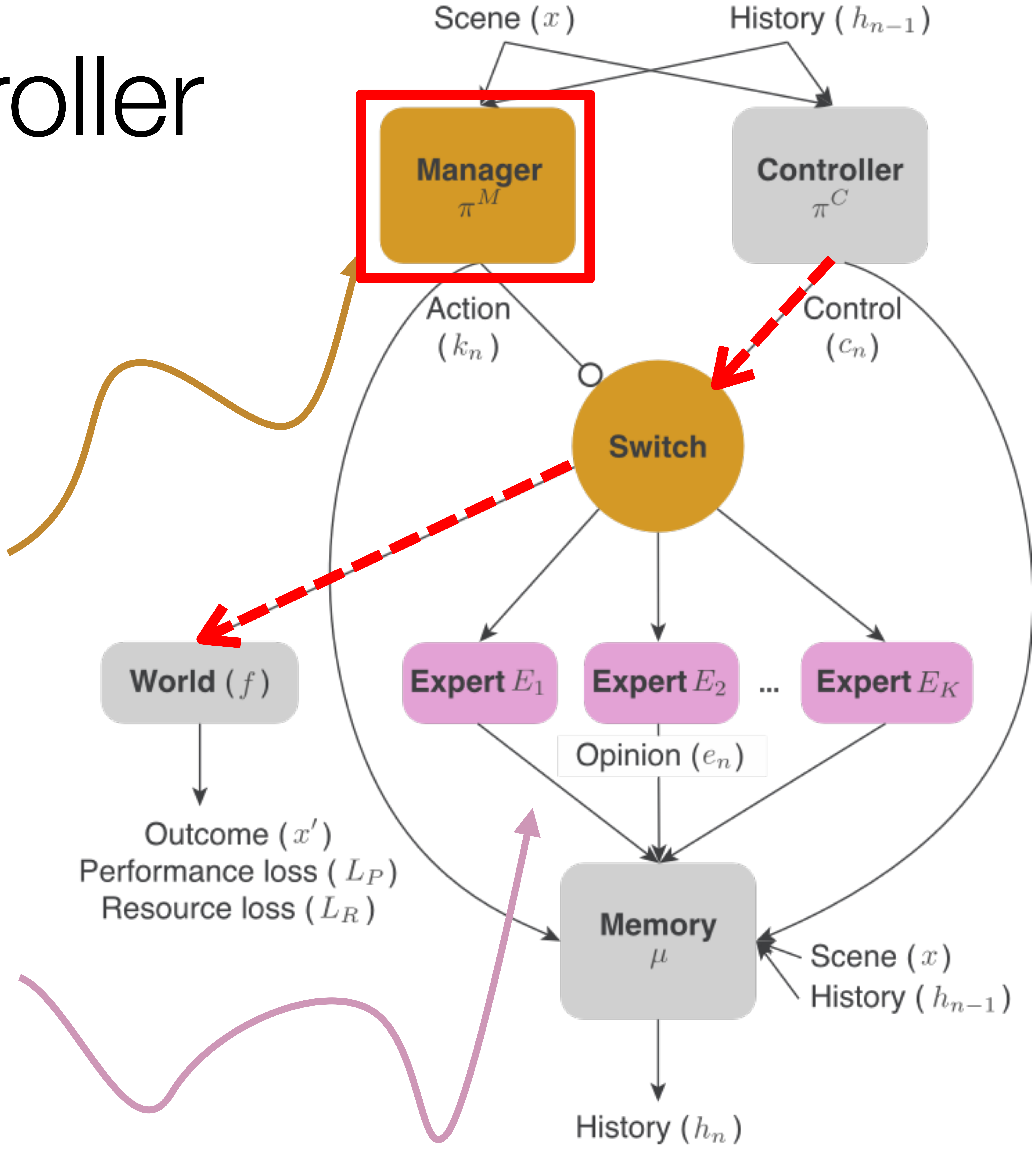
Experts**: Different models of the world, each with **different resource costs



Metacontroller

Manager (MLP)**: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to **minimize performance loss + resource loss

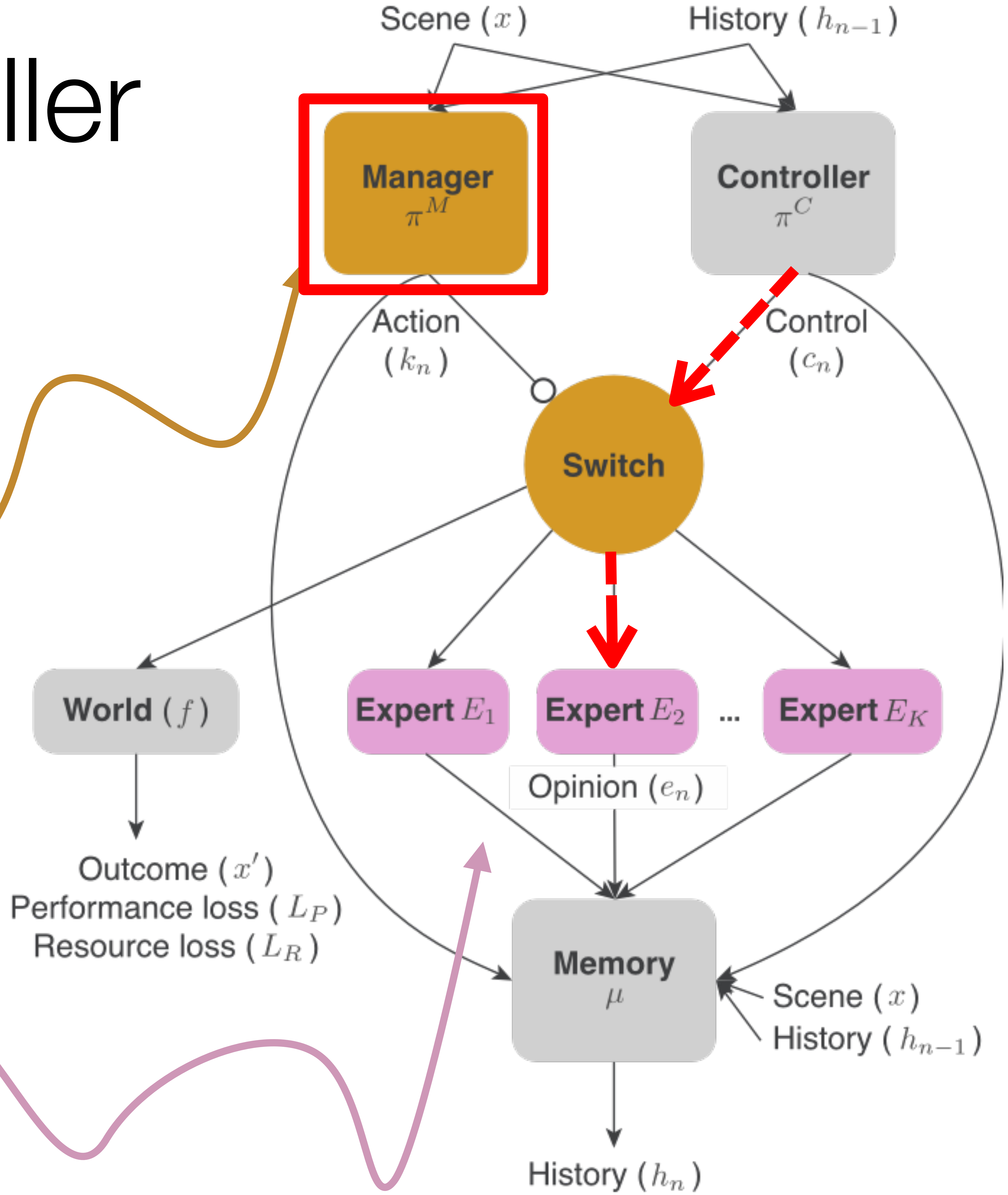
Experts**: Different models of the world, each with **different resource costs



Metacontroller

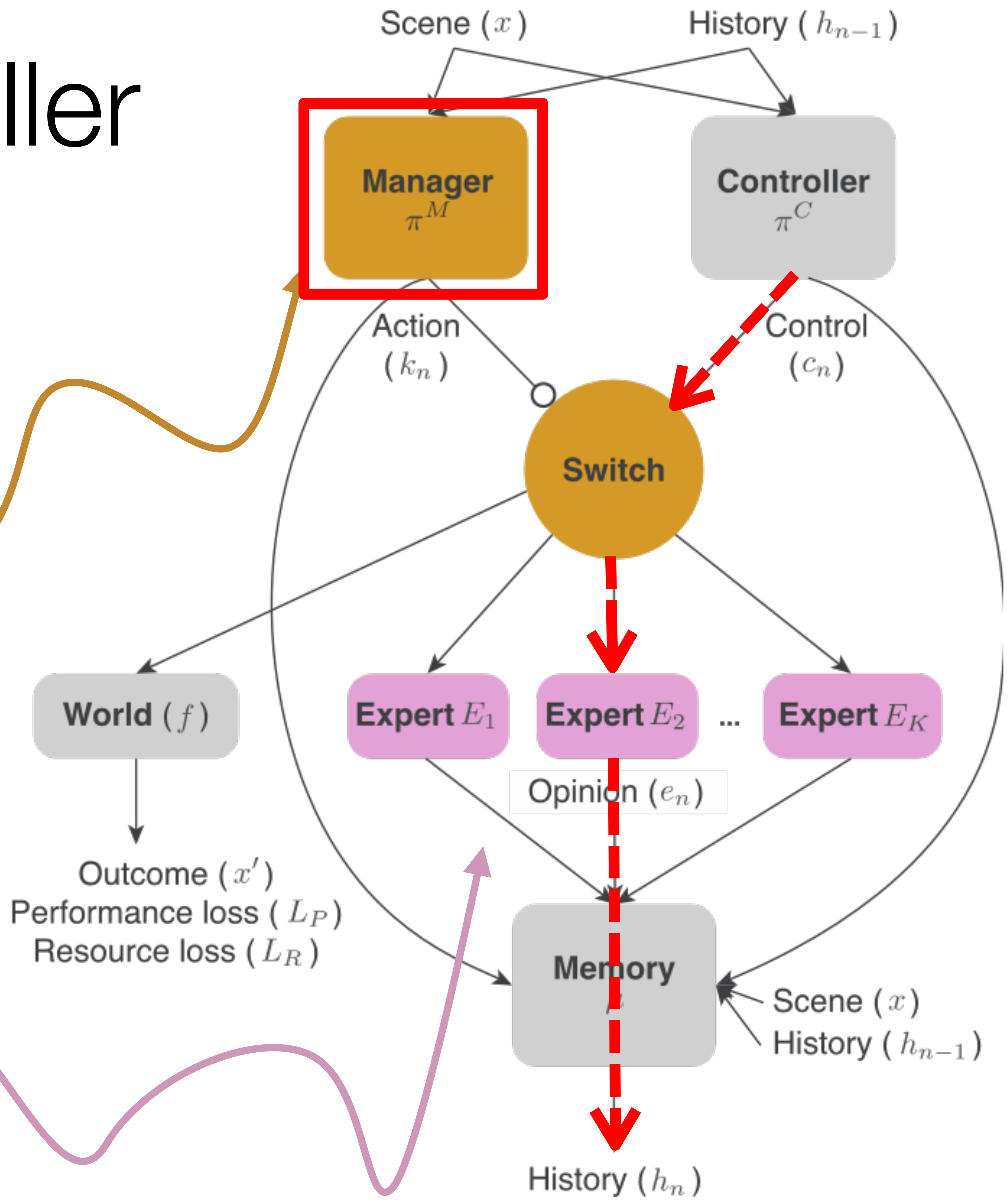
__Manager (MLP)__: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to __minimize performance loss + resource loss__

__Experts__: Different models of the world, each with __different resource costs__



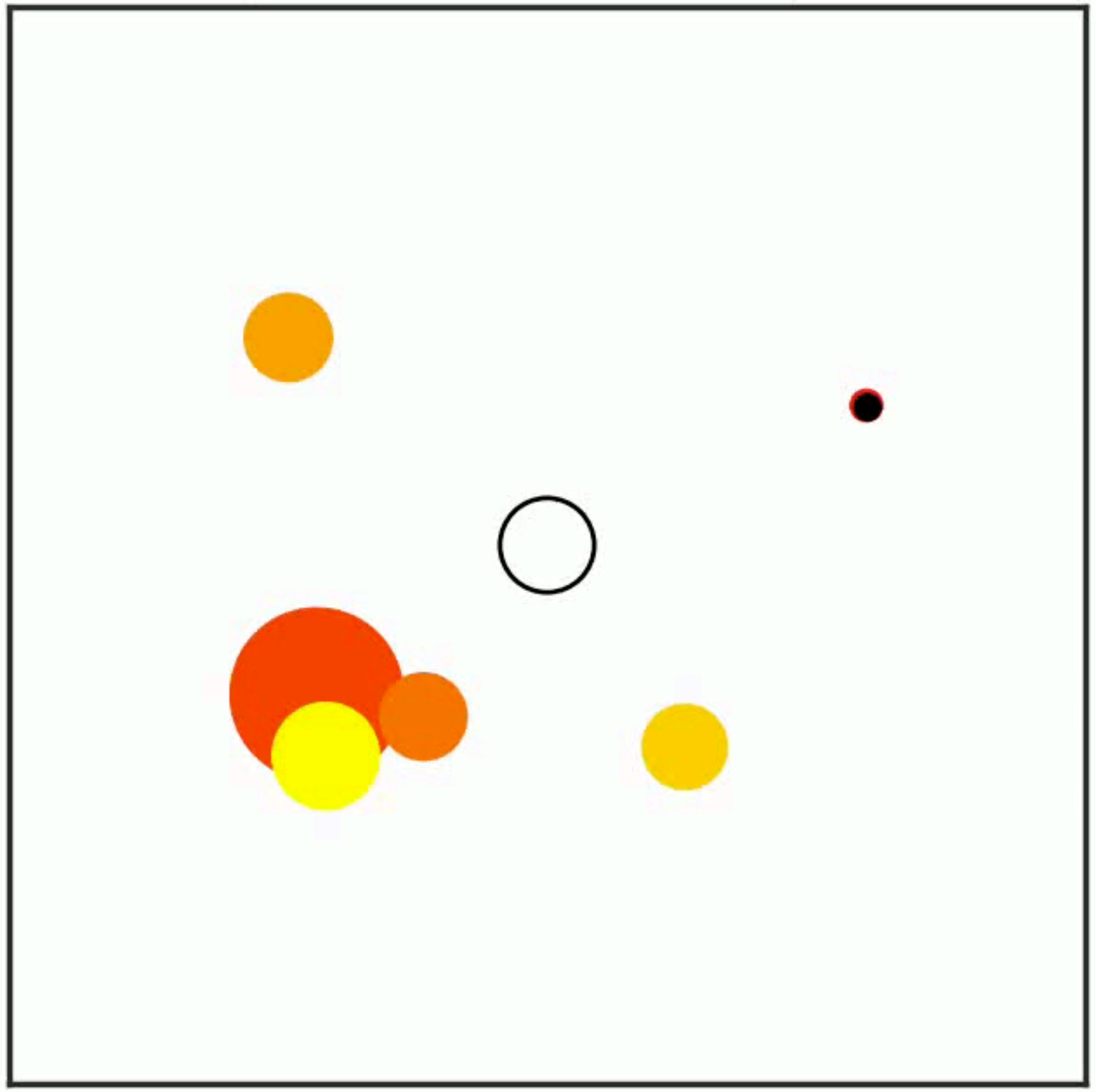
Metacontroller

Manager (MLP)*: Meta-level policy that determines whether to send the proposed control to the world, or to an expert, in order to* ***minimize performance loss + resource loss***

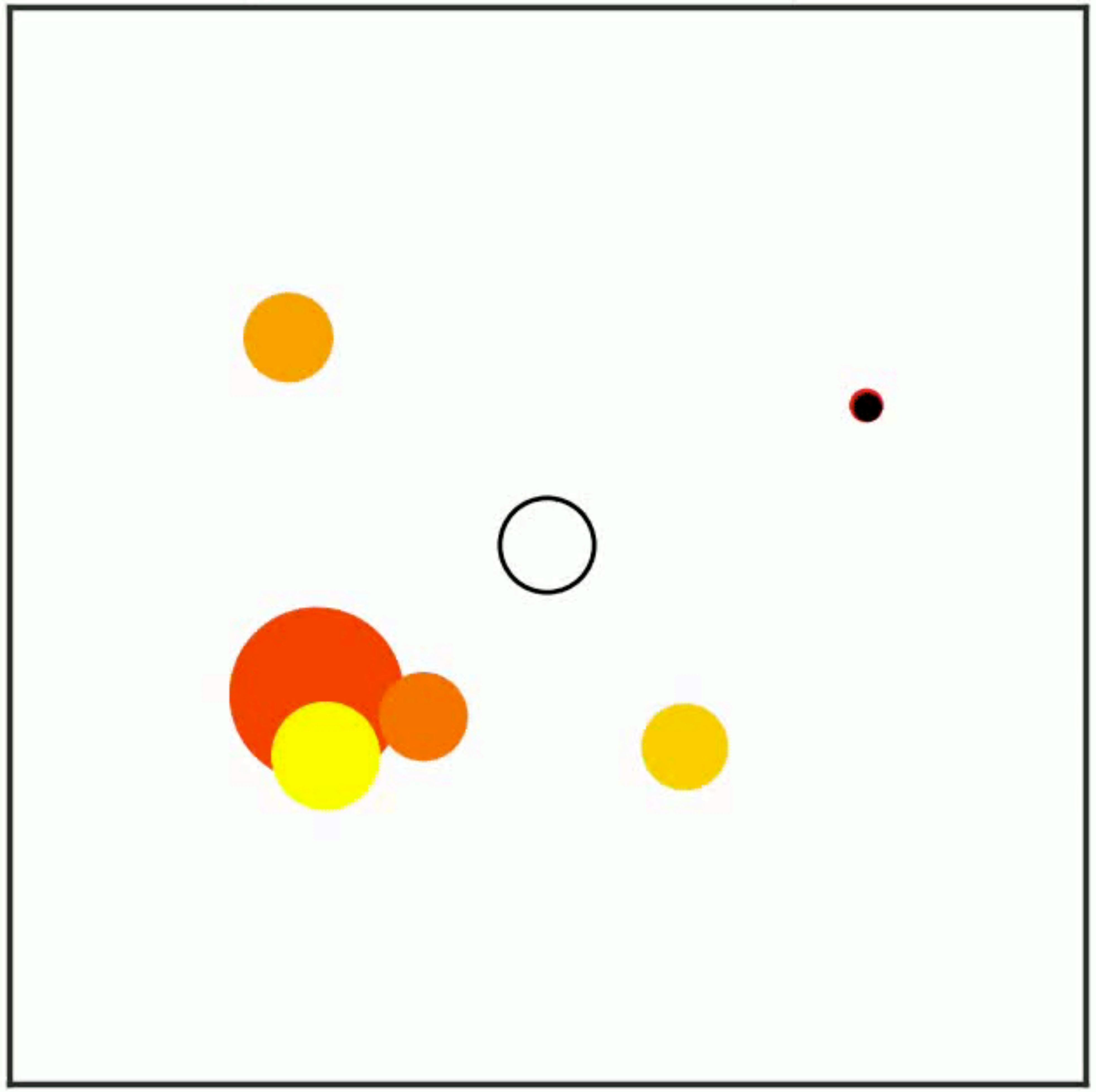


Experts*:*
Different models of the world, each with ***different resource costs***

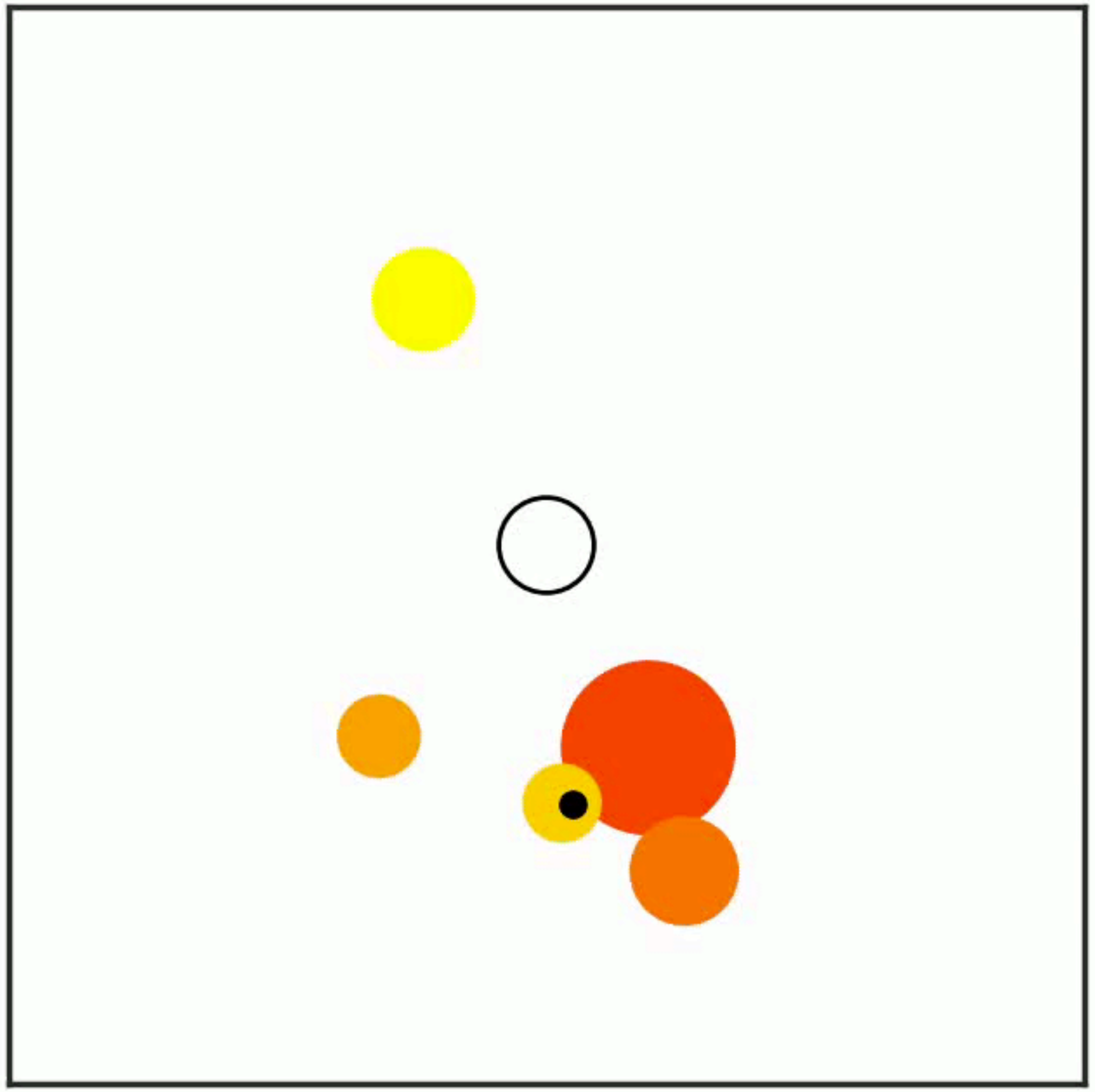
Metacontroller: simple task



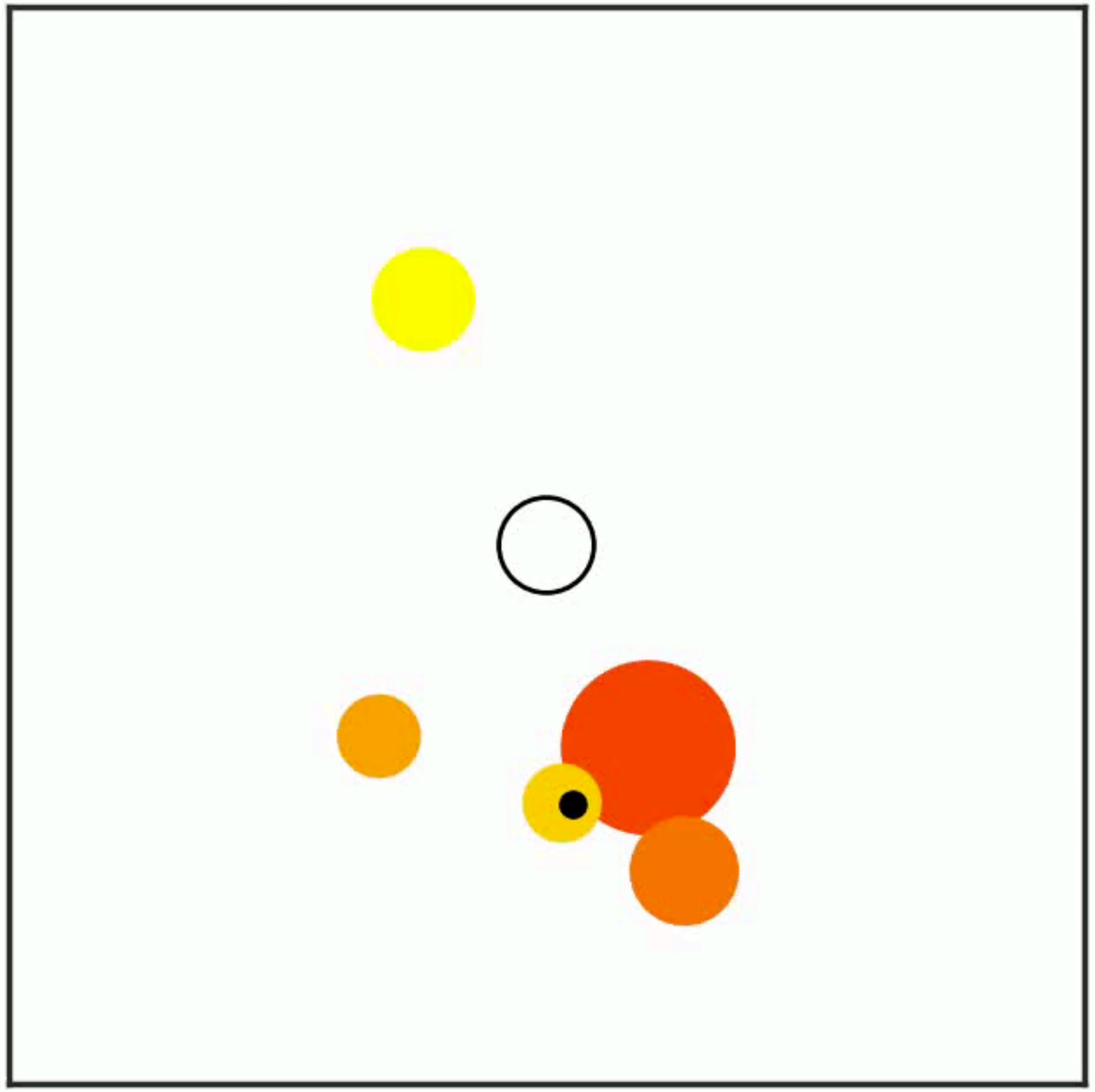
Metacontroller: simple task



Metacontroller: hard task

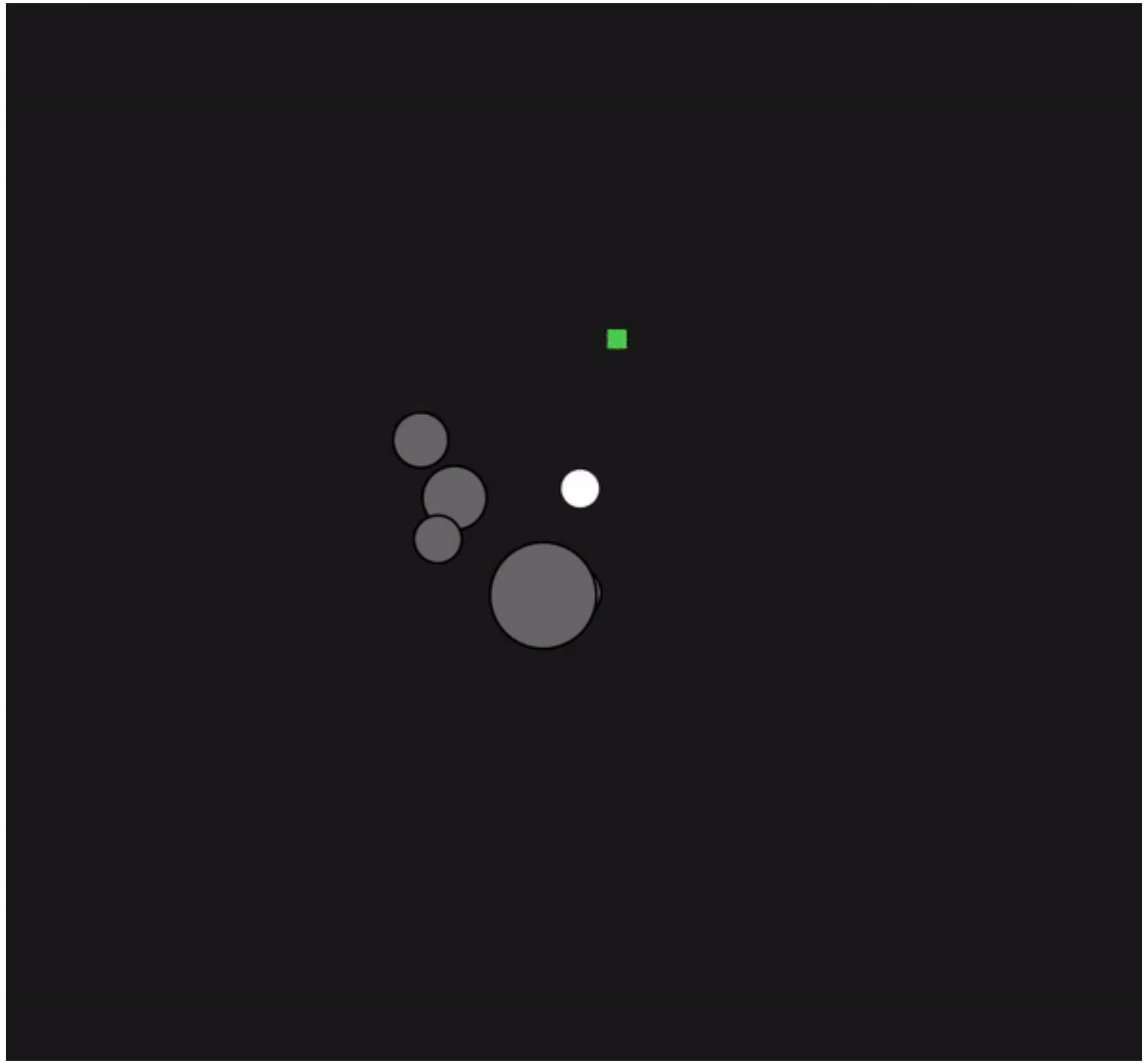


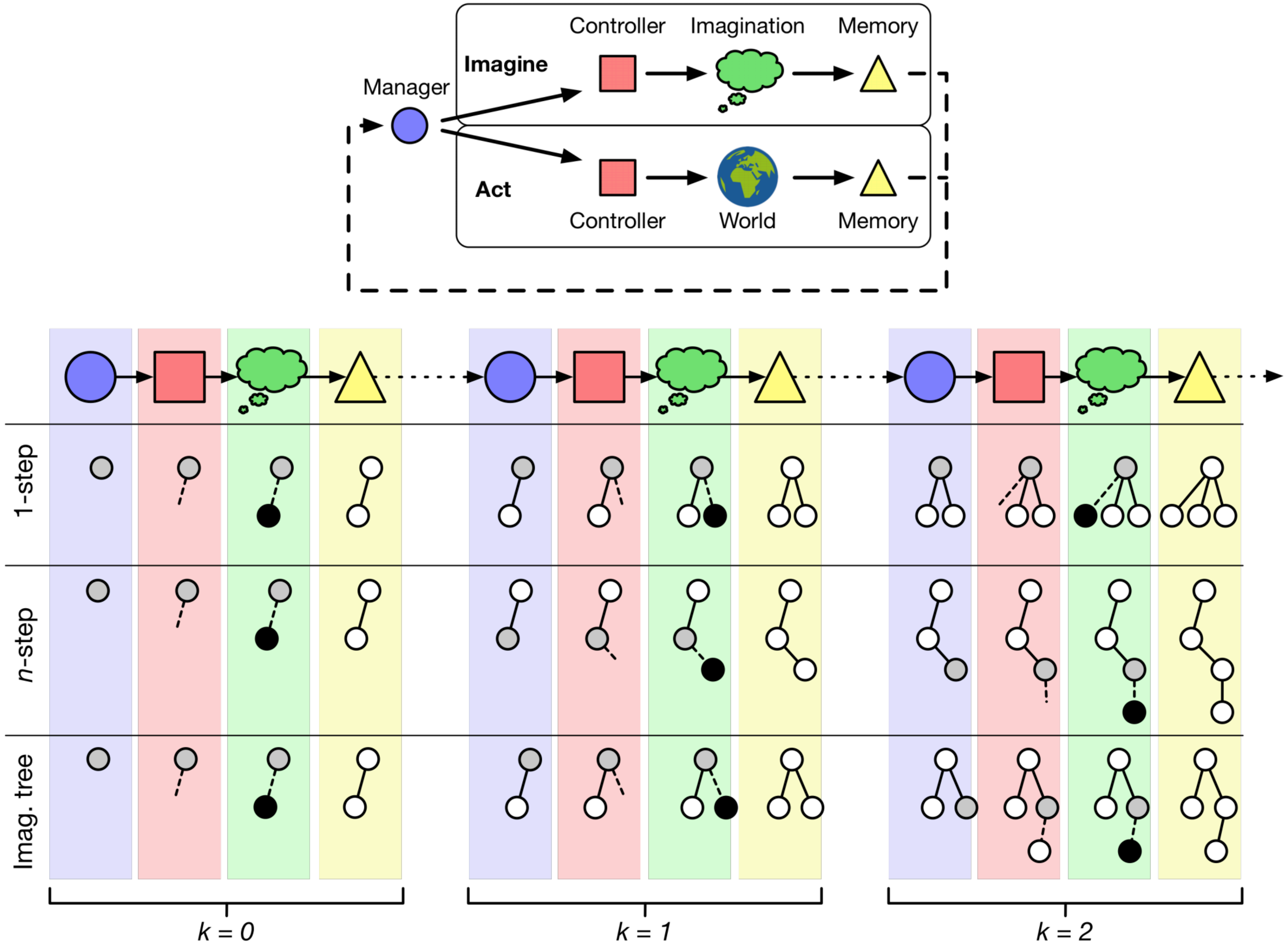
Metacontroller: hard task



Imagination-Based Planner

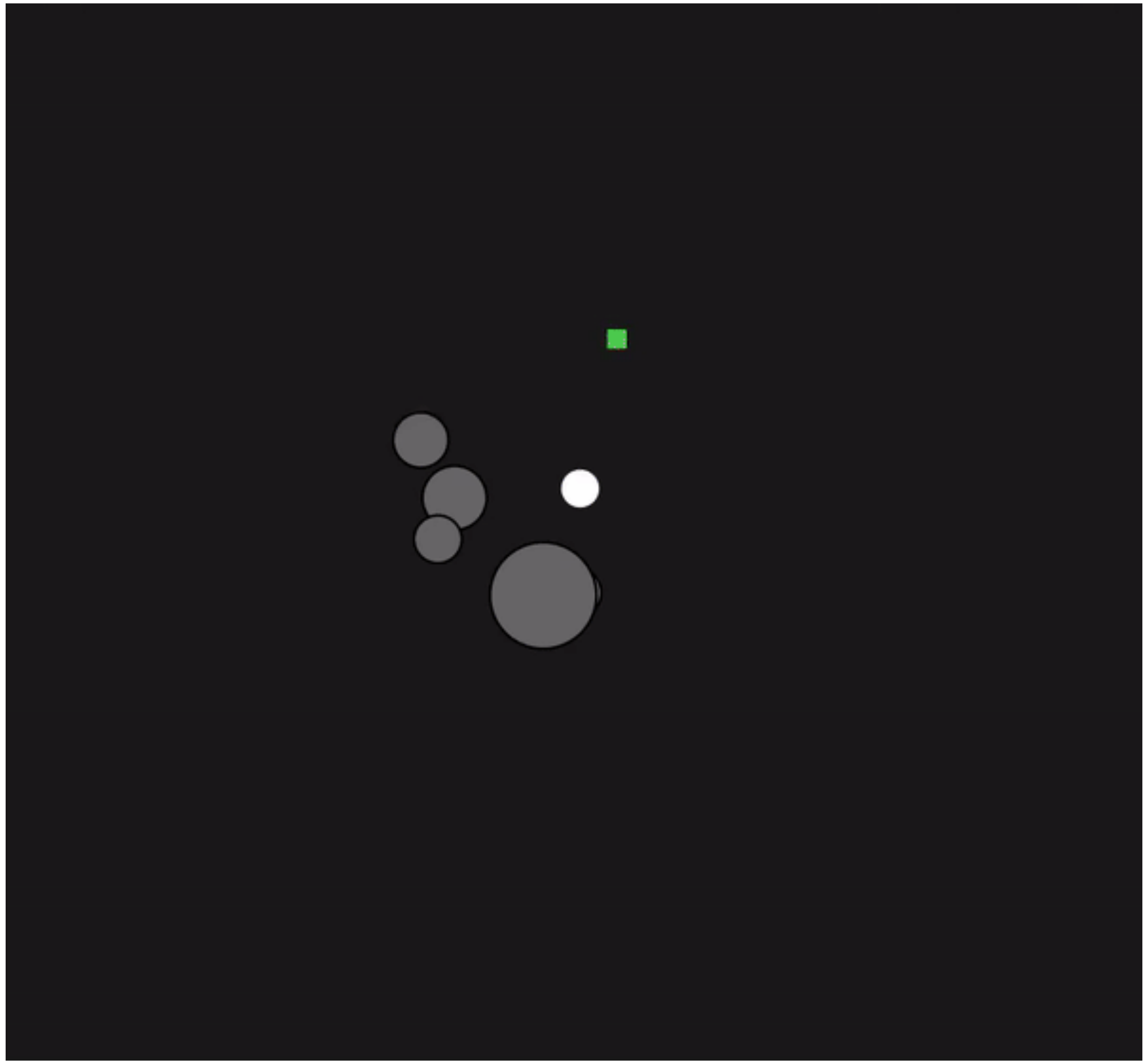
Pascanu, Li, et al. (2017)

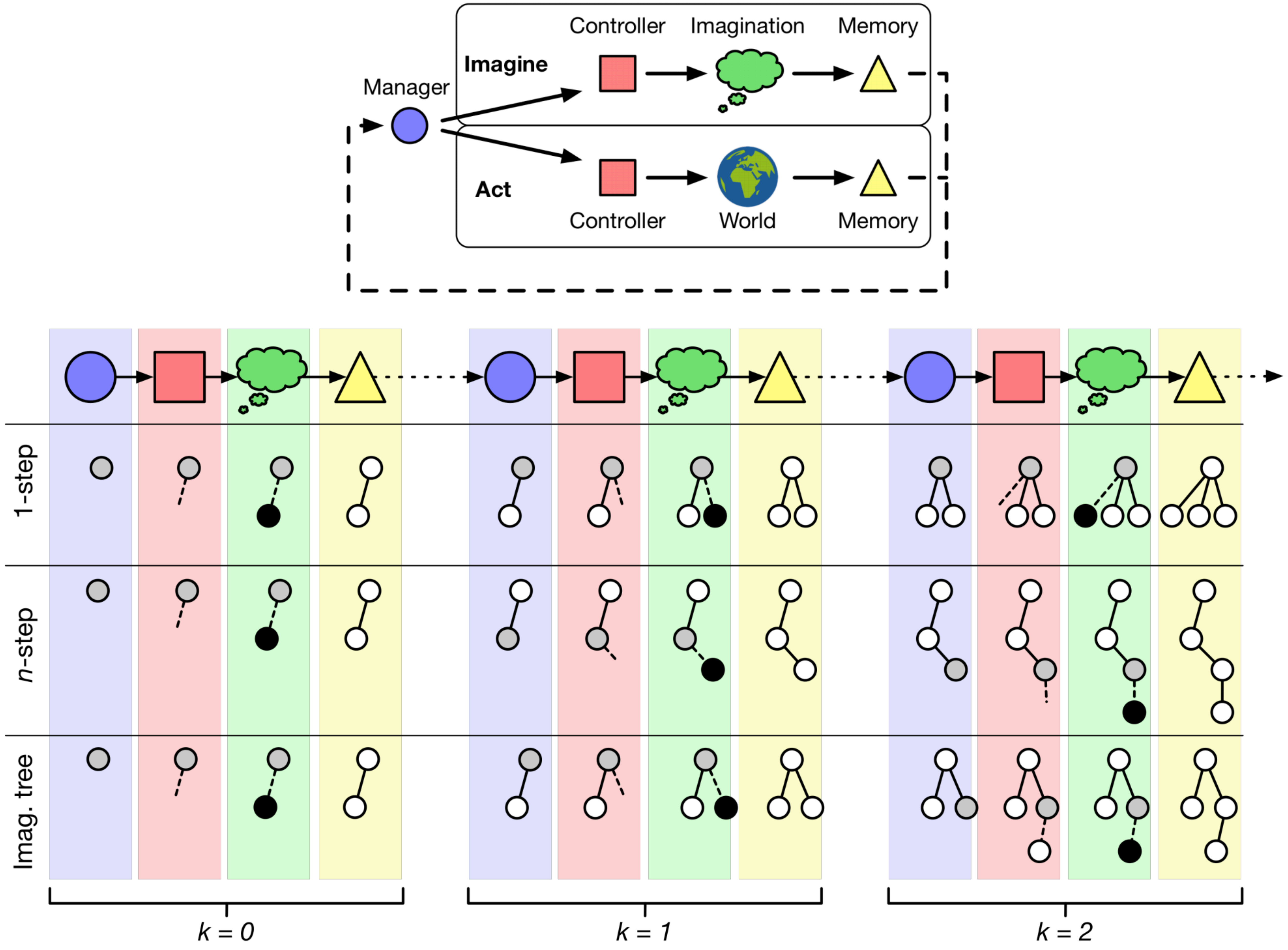




Imagination-Based Planner

Pascanu, Li, et al. (2017)

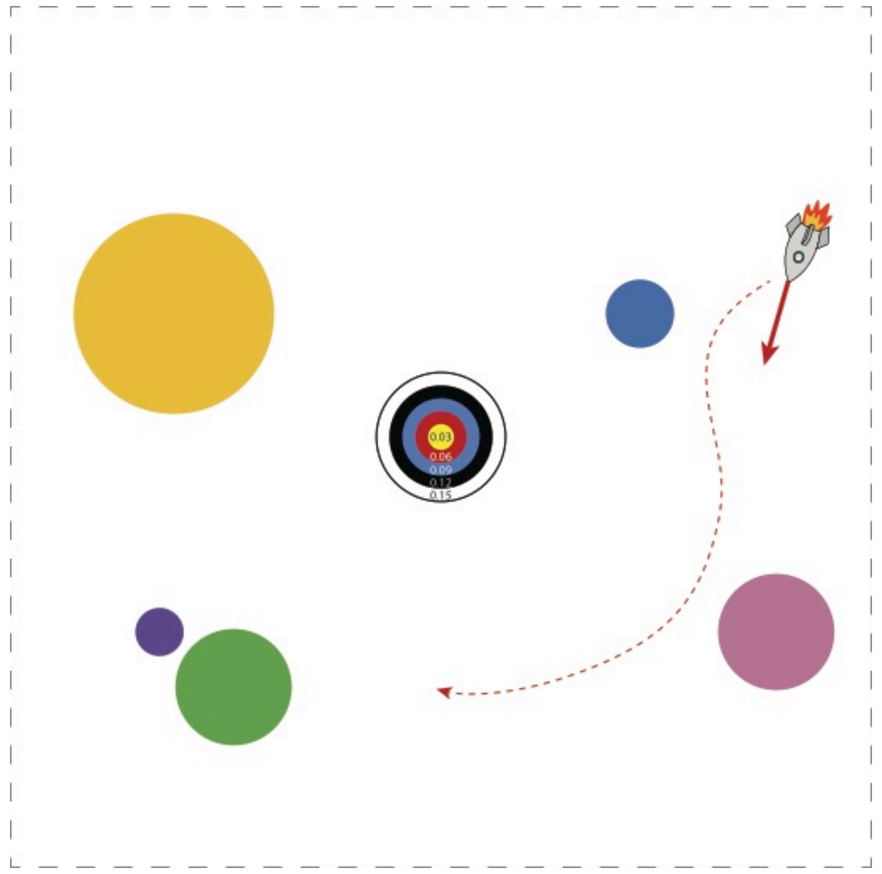




Outline

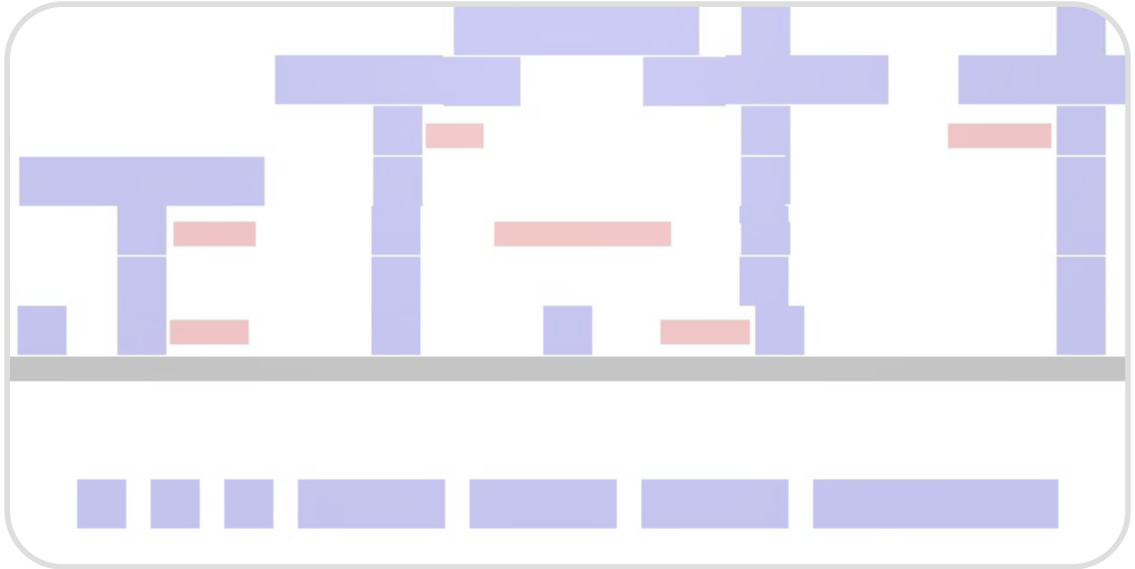
1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

   Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (2017, ICLR)



2. **Structured Representation**: the format of the data that computations are performed over.

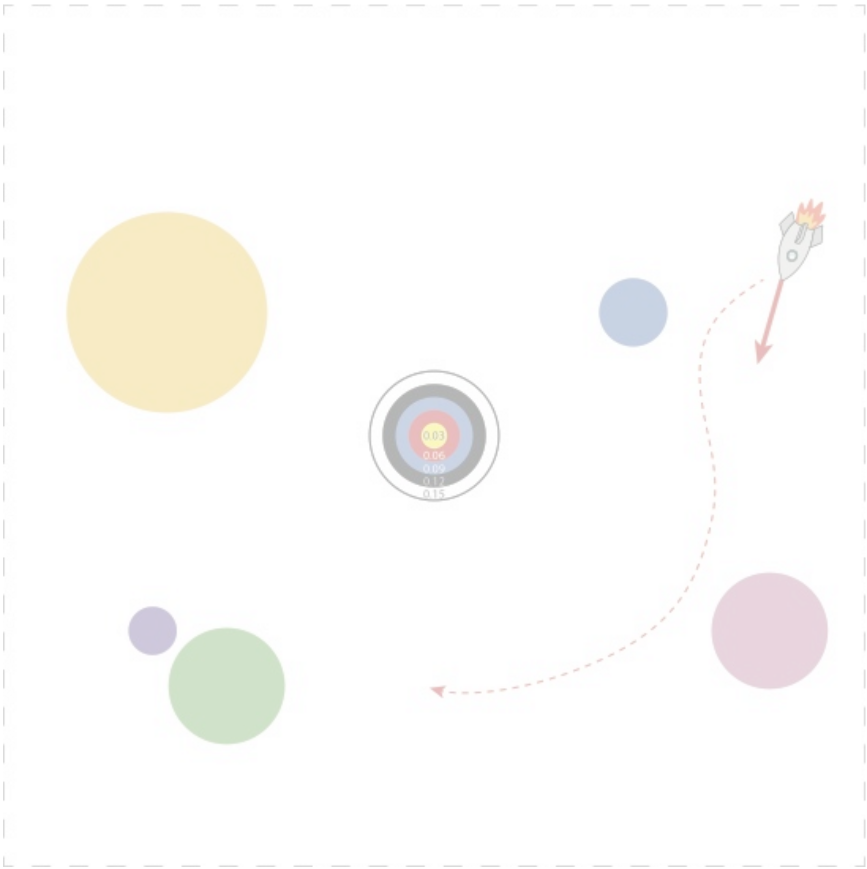
   Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018, arXiv)
   Bapst, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia, & Hamrick (2019, ICML)*



Outline

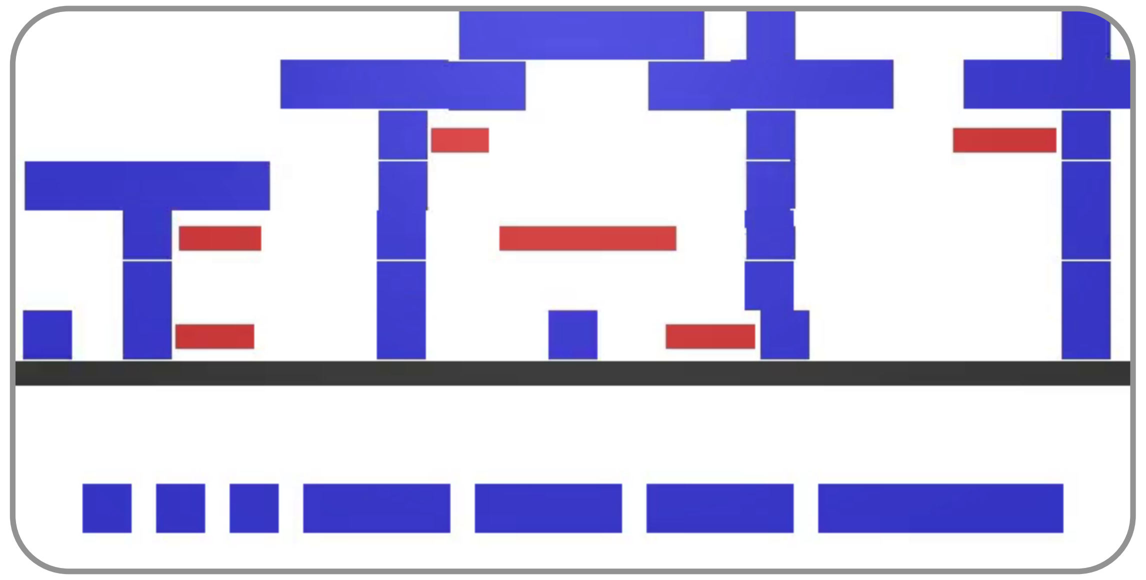
1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

   Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (2017, ICLR)



2. **Structured Representation**: the format of the data that computations are performed over.

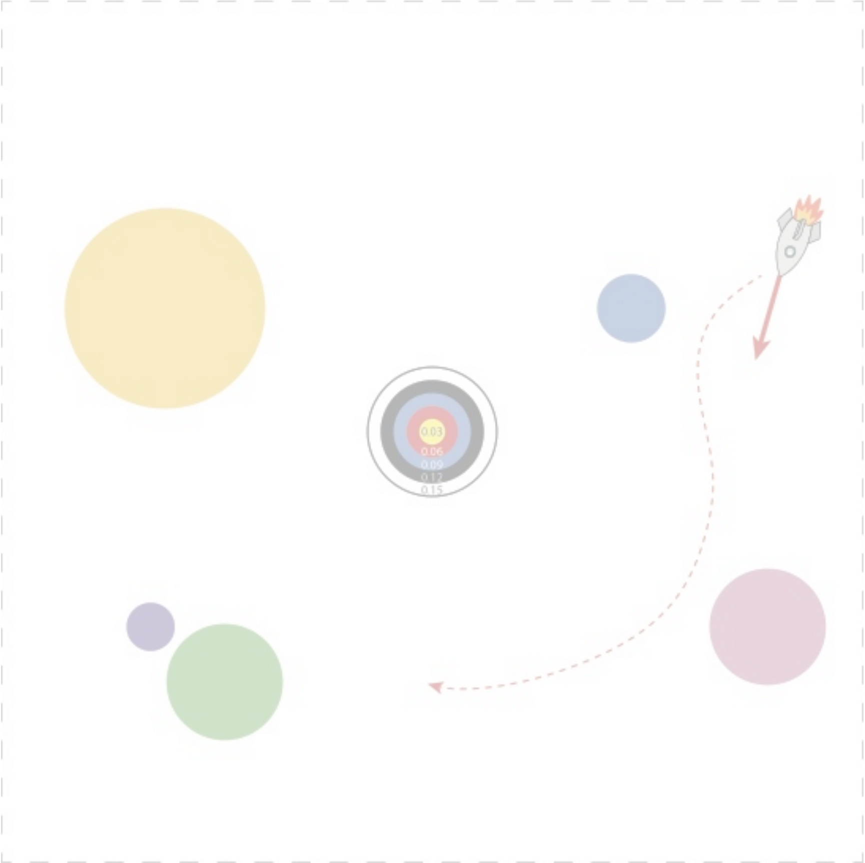
   Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018, arXiv)
   Bapst, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia, & Hamrick (2019, ICML)*



Outline

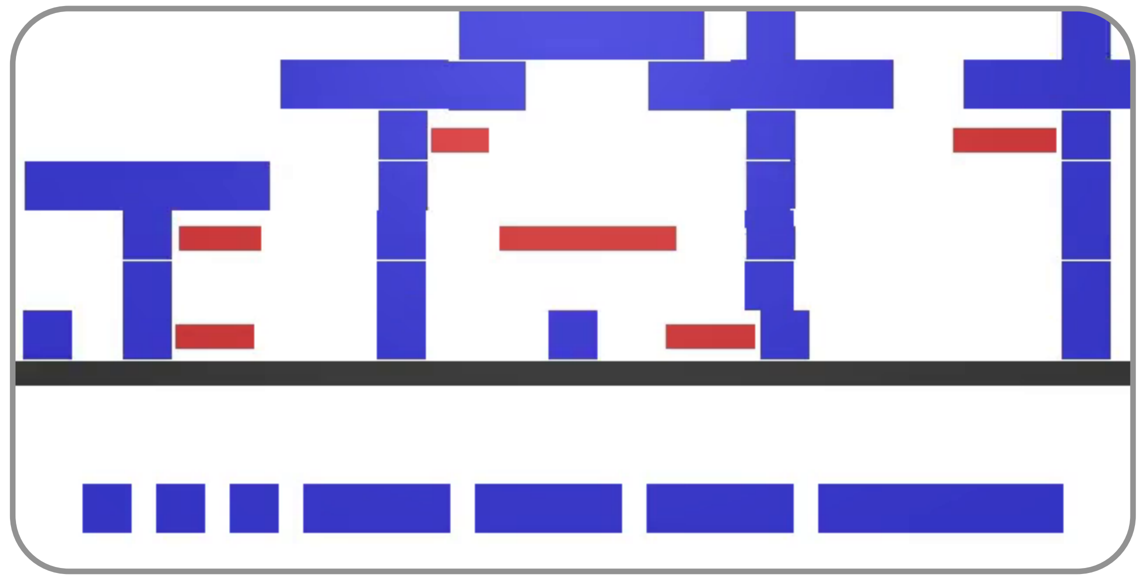
1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

   Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (2017, ICLR)



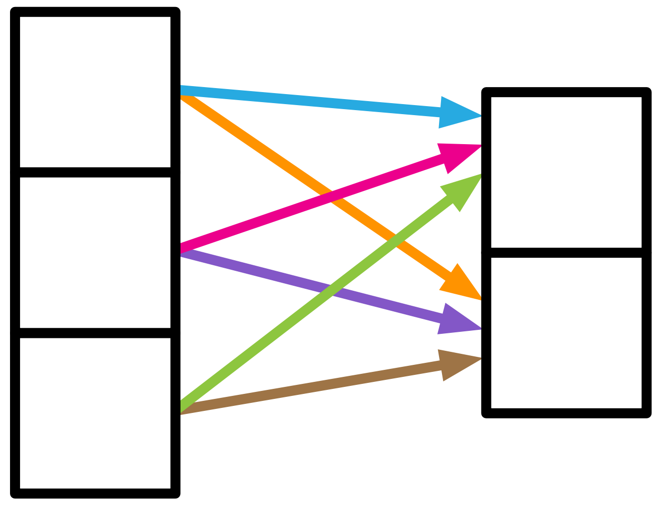
2. **Structured Representation**: the format of the data that computations are performed over.

   Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018, arXiv)
   Bapst, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia, & Hamrick (2019, ICML)*

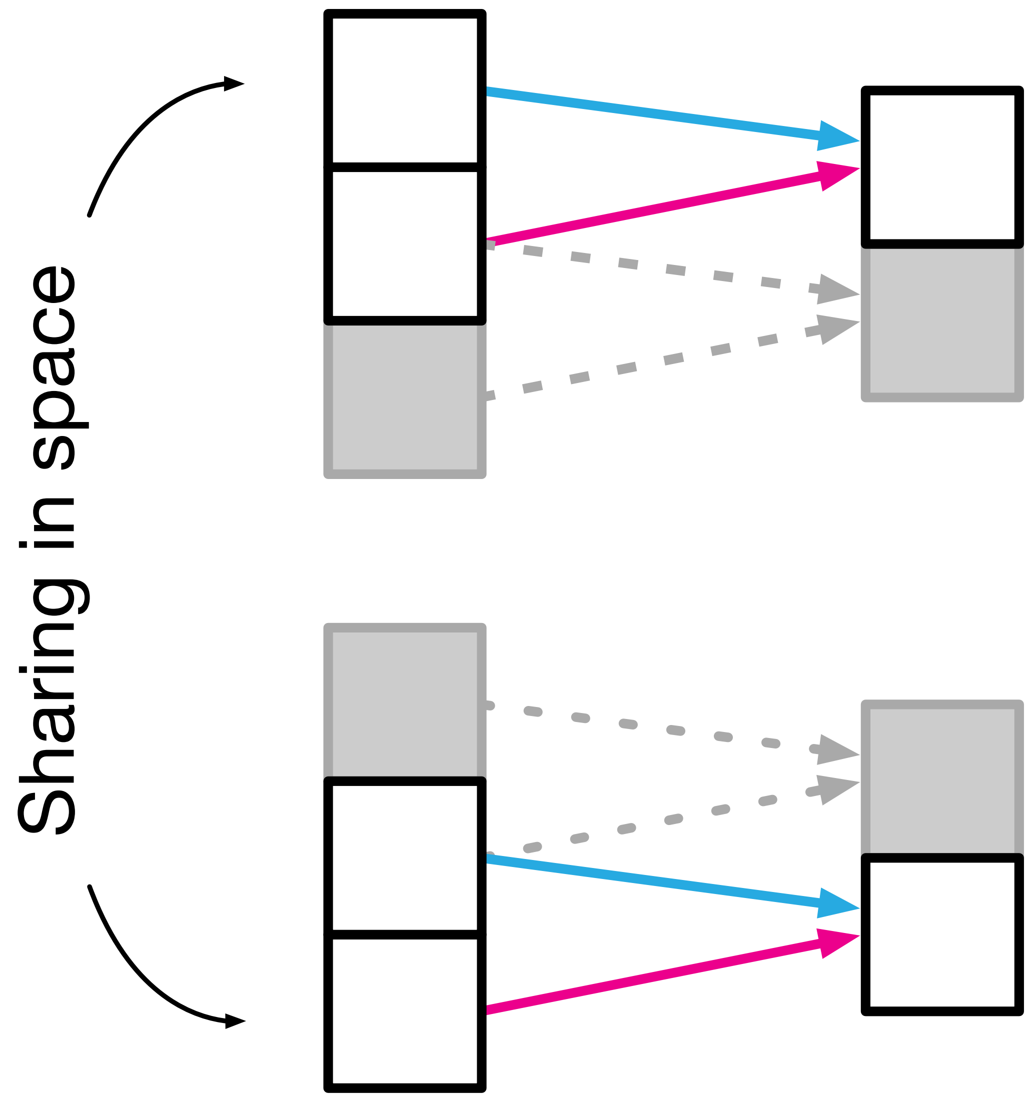


Led by Victor Bapst and
Alvaro Sanchez-Gonzalez

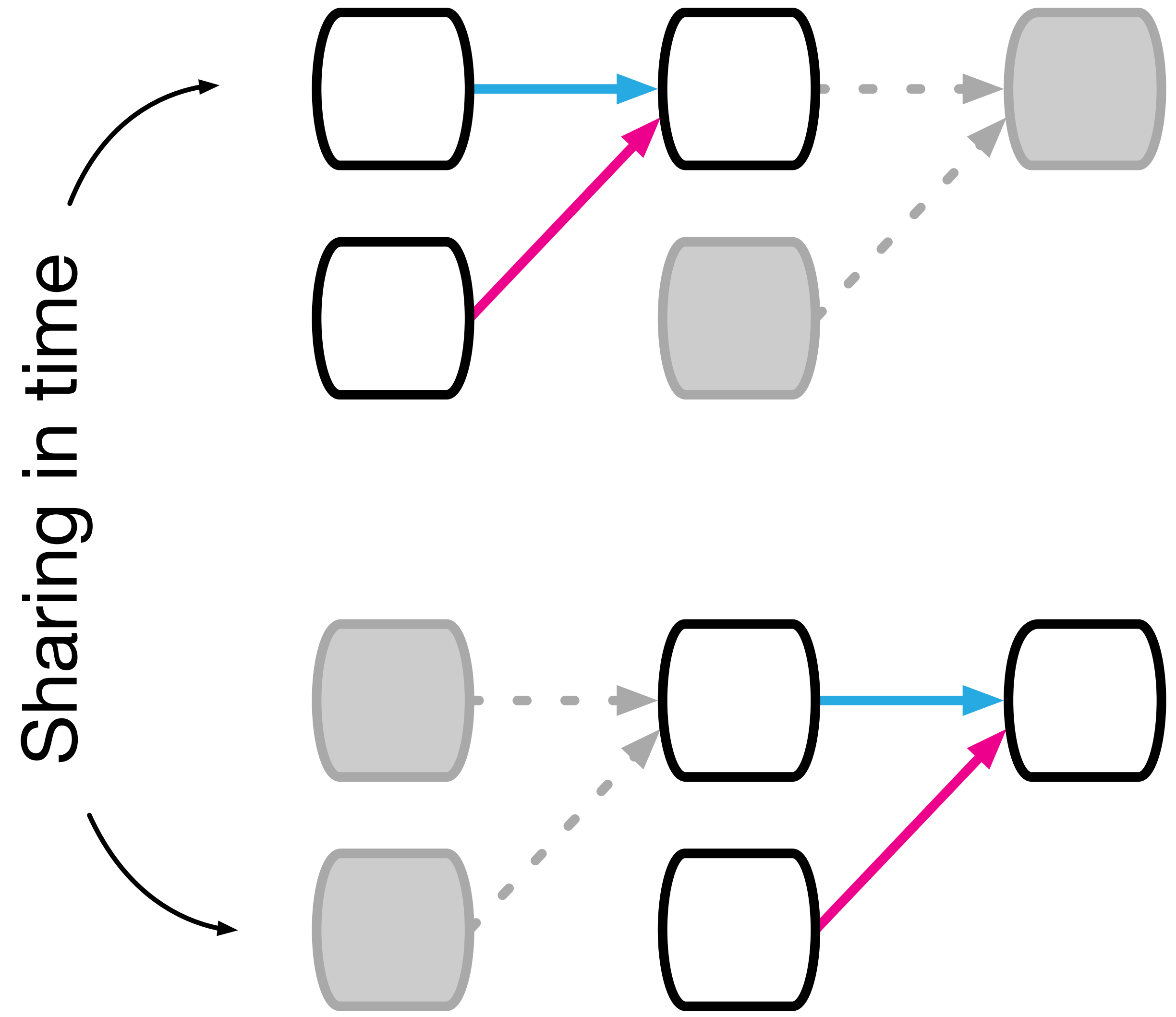
Structure in Deep Networks



Fully-Connected Layer
Unstructured computation
Unstructured representation



Convolutional Layer
Structured computation
Unstructured representation



Recurrent Layer
Structured computation
Unstructured representation

Graph Networks

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

1. Take **graphs** as input, return graphs as output

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

1. Take **graphs** as input, return graphs as output

2. Invariant to the **permutation** of the nodes and edges

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

1. Take **graphs** as input, return graphs as output

2. Invariant to the **permutation** of the nodes and edges

3. Scale to different **numbers** of nodes and edges

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

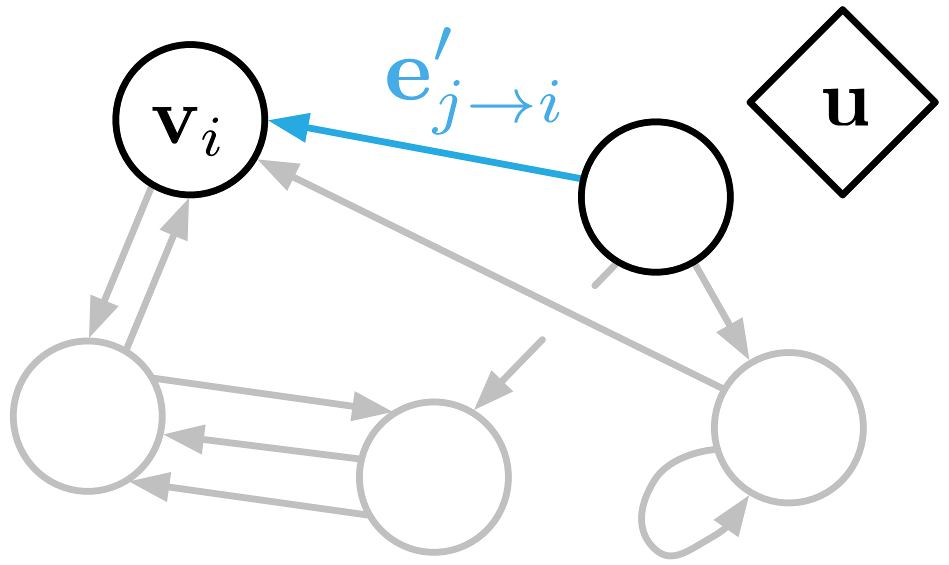
Edges	**Nodes**	**Globals**
E	V	$\mathbf{u}$

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

Edges	Nodes	Globals
E	V	$\mathbf{u}$

Edge update



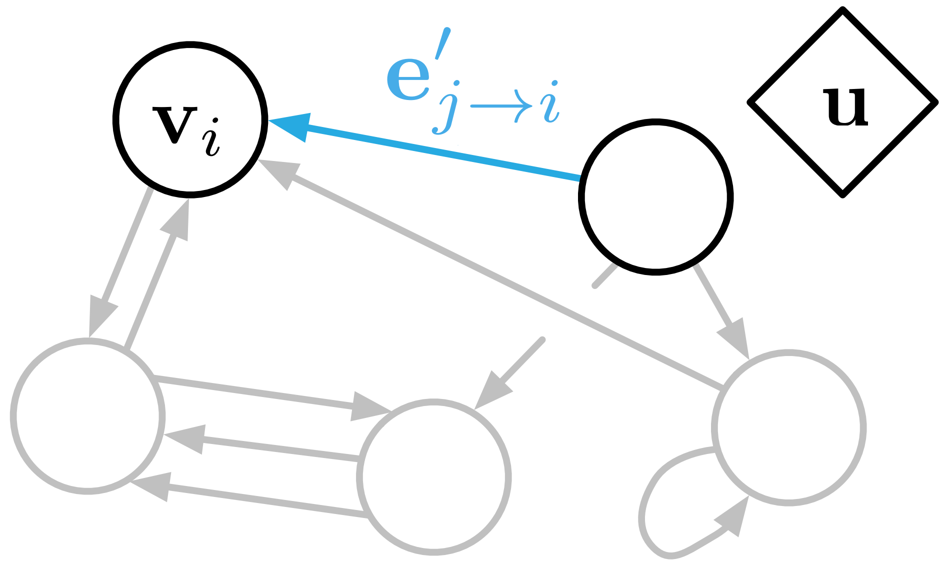
$$\mathbf{e}'_{i\to j} = \phi_e(\mathbf{v}_i, \mathbf{v}_j, \mathbf{e}_{i\to j}, \mathbf{u})$$

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

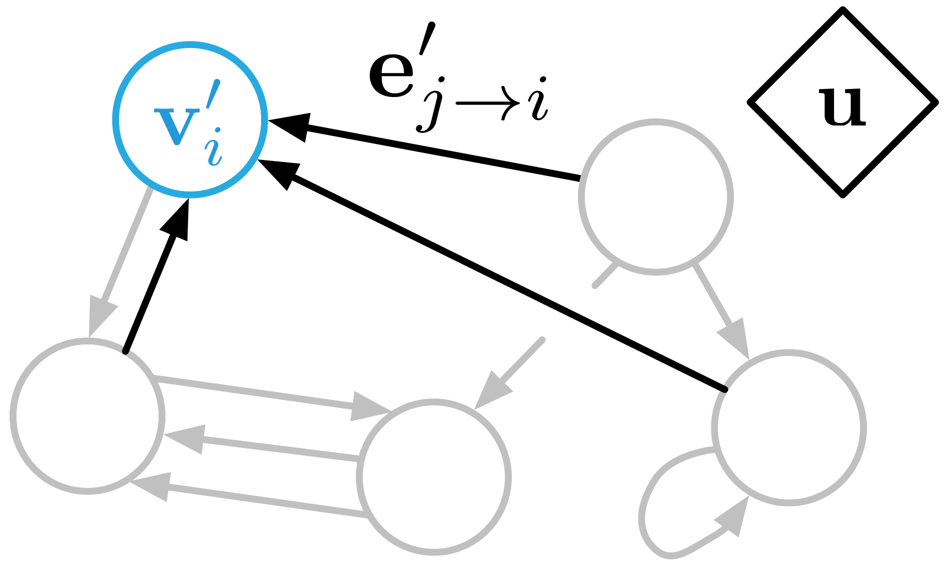
Edges	Nodes	Globals
E	V	$\mathbf{u}$

Edge update



$$\mathbf{e}'_{i\to j} = \phi_e(\mathbf{v}_i, \mathbf{v}_j, \mathbf{e}_{i\to j}, \mathbf{u})$$

Node update



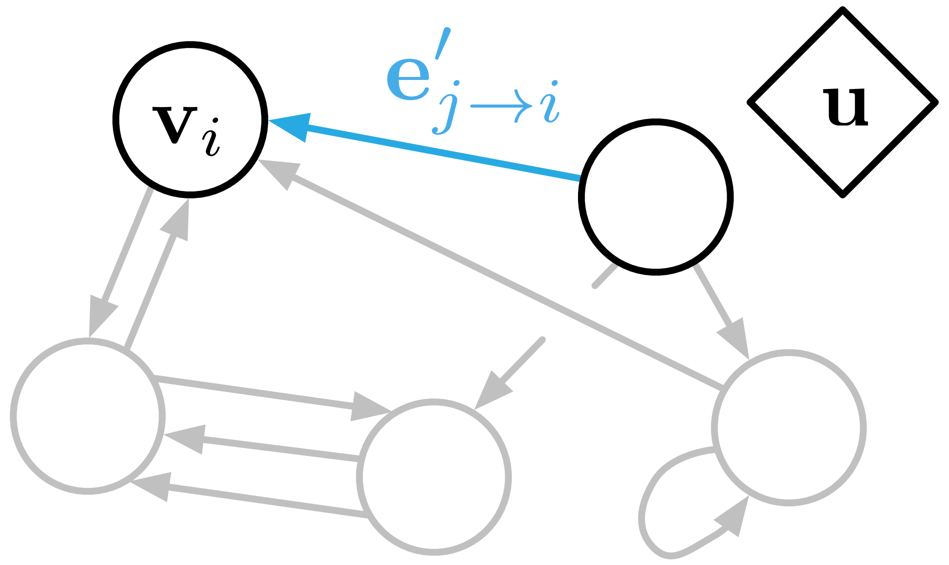
$$\mathbf{v}'_i = \phi_v(\mathbf{v}_i, \sum_j \mathbf{e}'_{j\to i}, \mathbf{u})$$

Graph Networks

Gori et al. (2005), Scarselli et al. (2005), Scarselli et al. (2009), Li et al. (2015), Gilmer et al. (2017)
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018)

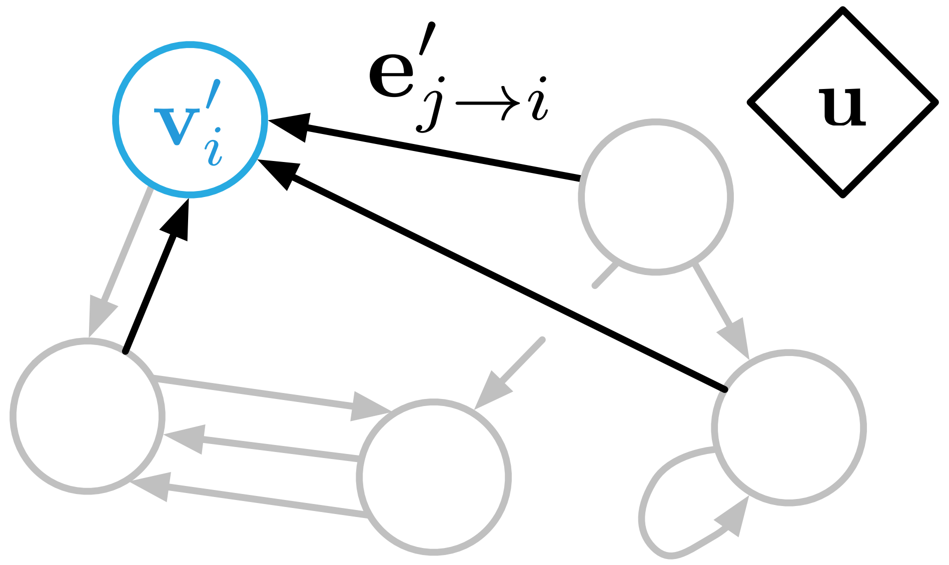
Edges	Nodes	Globals
E	V	$\mathbf{u}$

Edge update



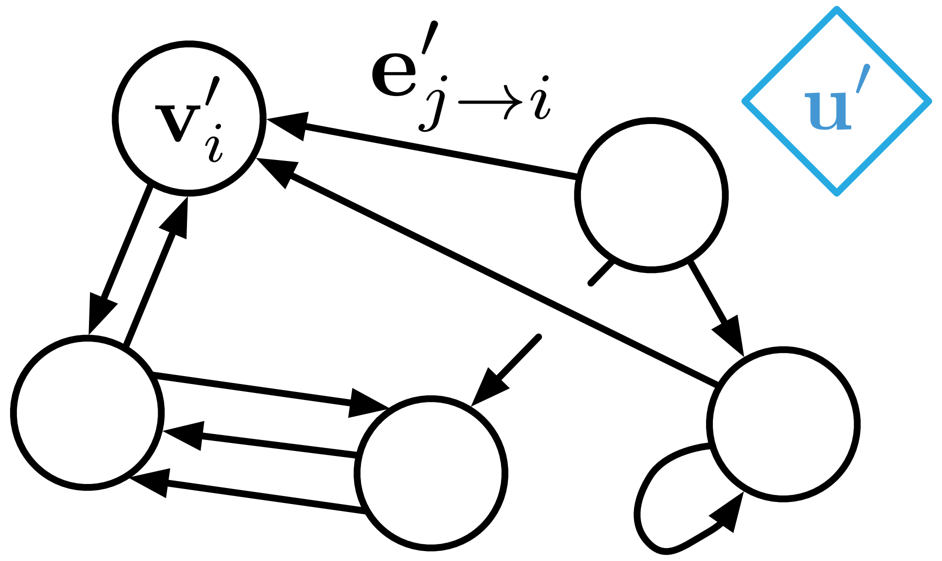
$$\mathbf{e}'_{i\to j} = \phi_e(\mathbf{v}_i, \mathbf{v}_j, \mathbf{e}_{i\to j}, \mathbf{u})$$

Node update



$$\mathbf{v}'_i = \phi_v(\mathbf{v}_i, \sum_j \mathbf{e}'_{j\to i}, \mathbf{u})$$

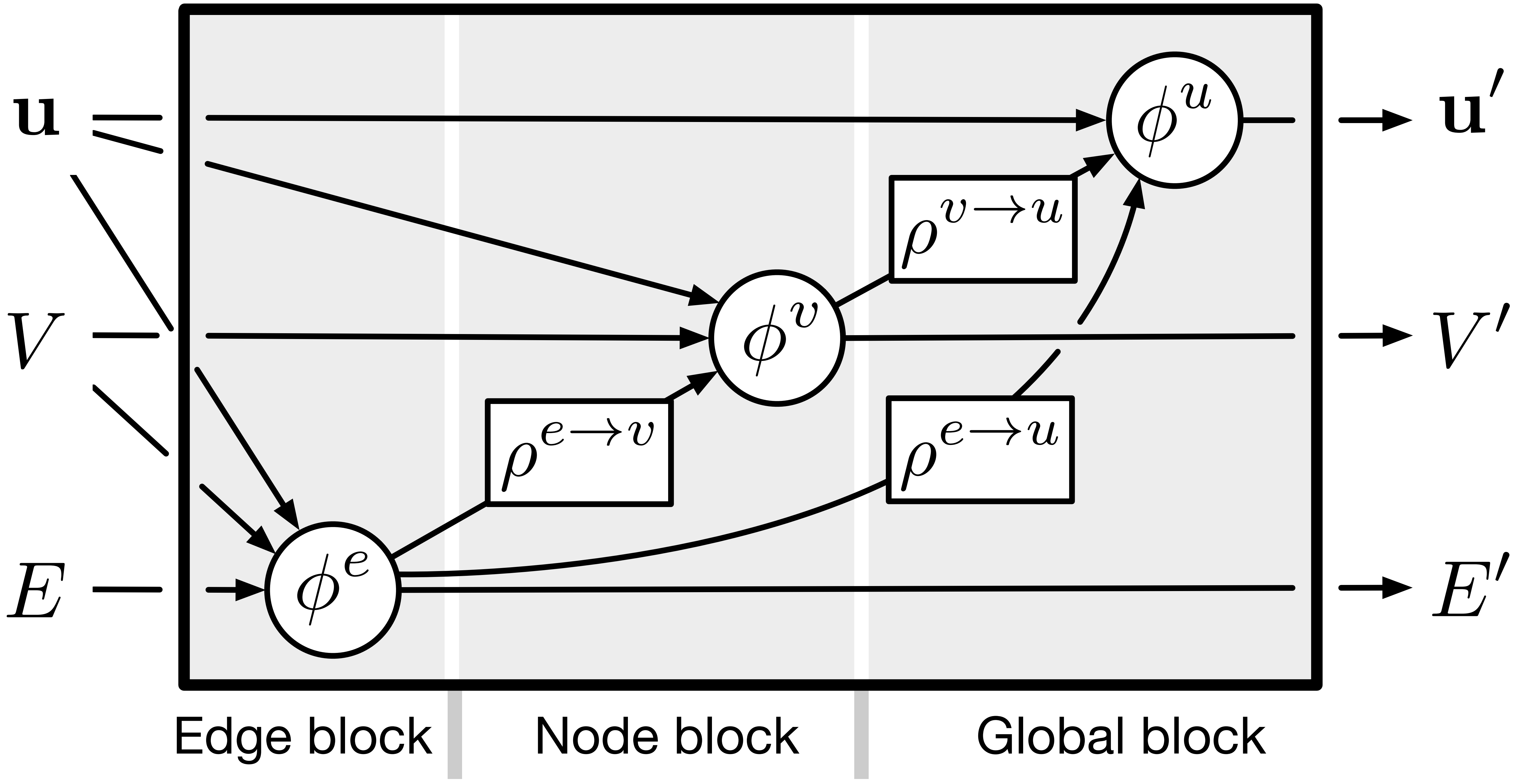
Globals update



$$\mathbf{u}' = \phi_u(\sum_i \mathbf{v}'_i, \sum_{i,j} \mathbf{e}'_{i\to j}, \mathbf{u})$$

Graph Networks

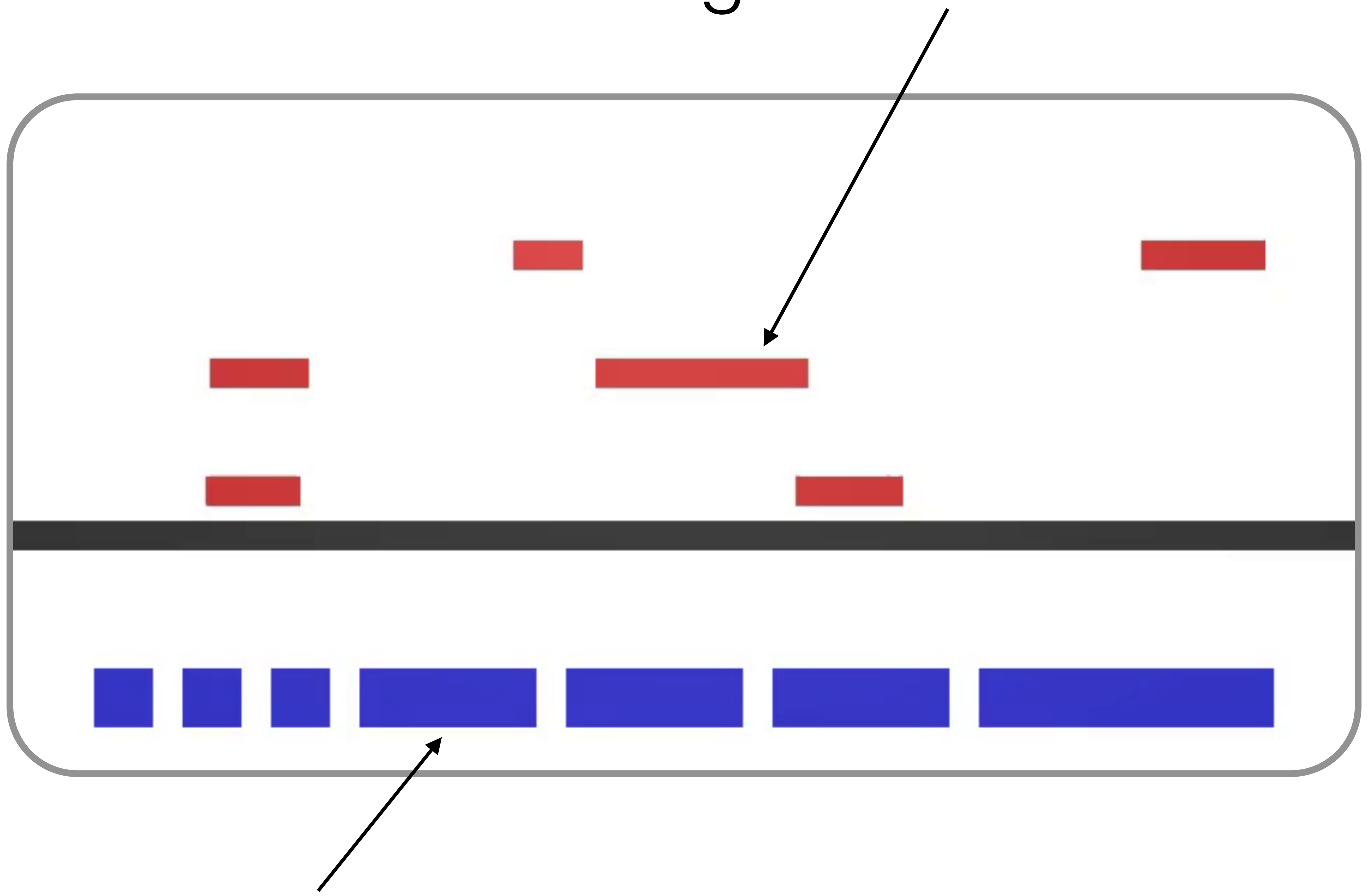
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (*arXiv* 2018)



Construction Tasks

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Avoid touching **obstacles**

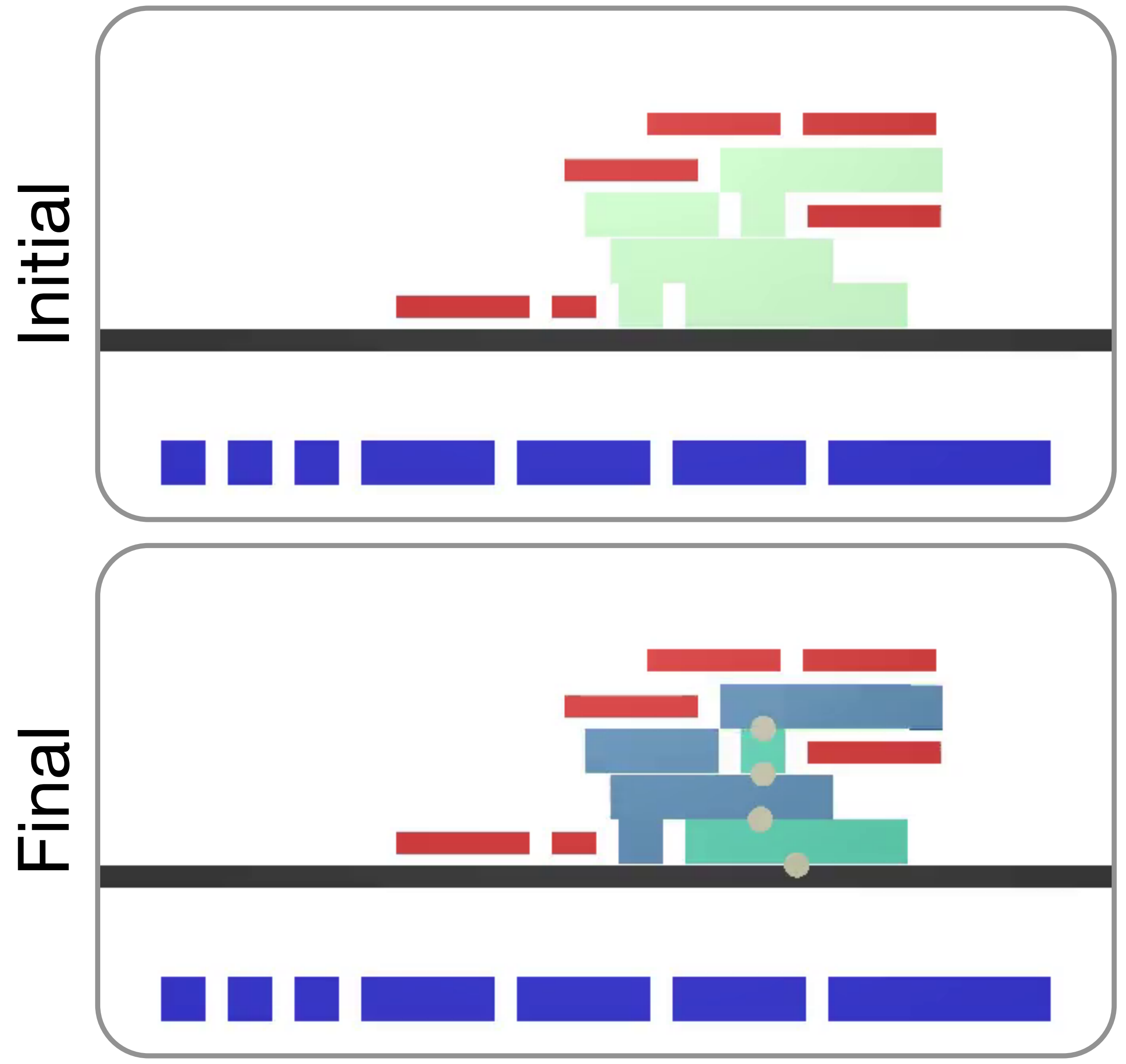


Pick up **blocks** and place them in the scene
(and optionally make them **sticky**)

Construction Tasks

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

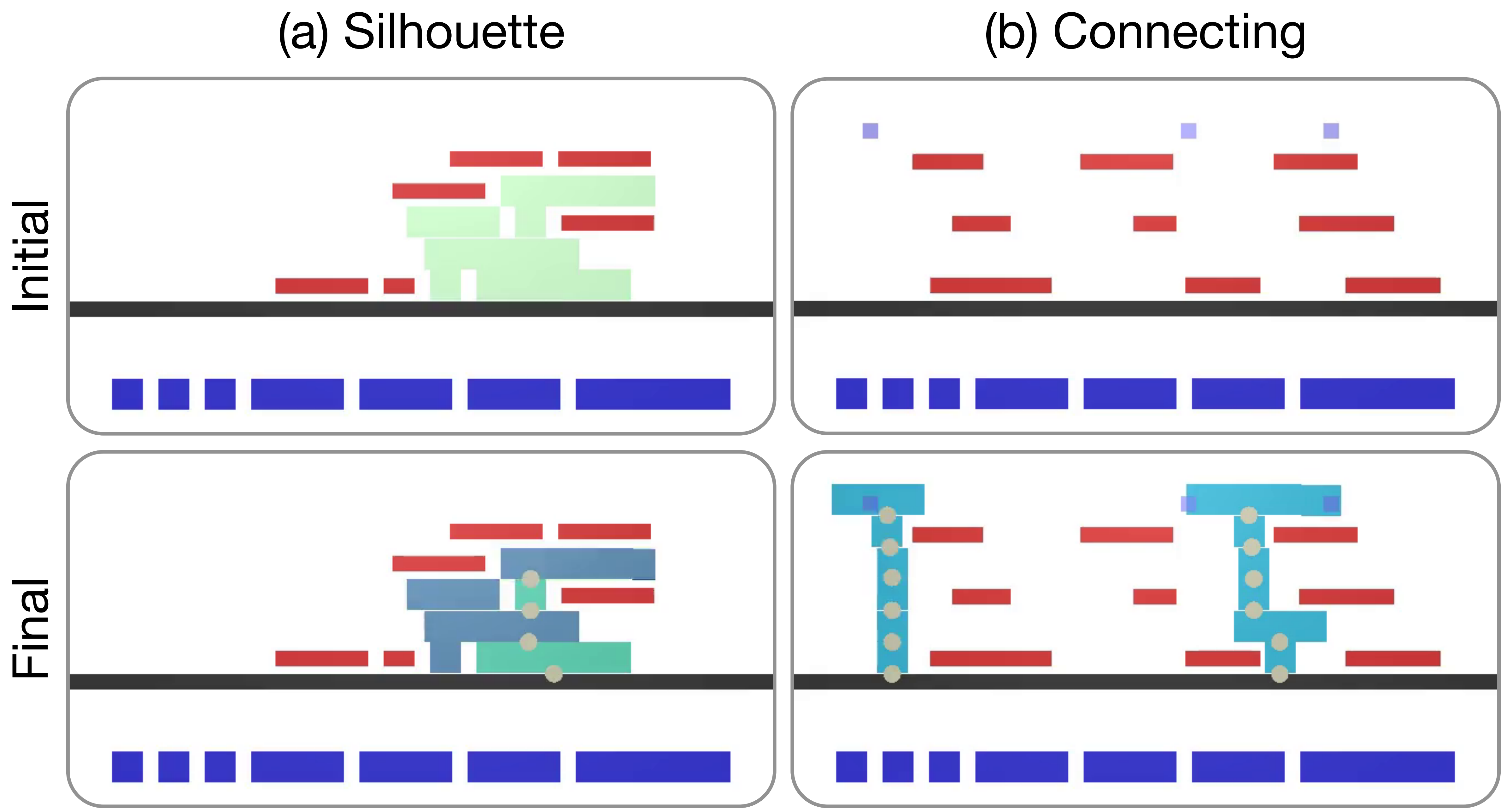
(a) Silhouette



+1 per target
-0.5 per sticky block

Construction Tasks

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

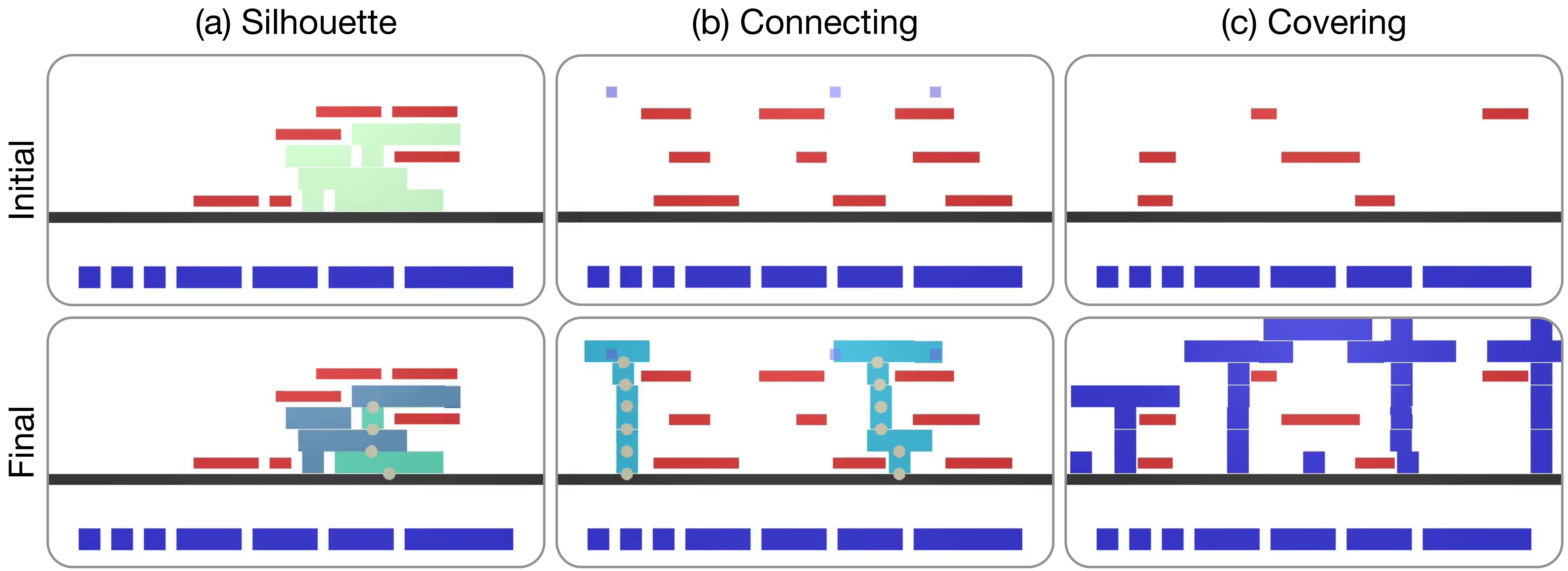


+1 per target
-0.5 per sticky block

+1 per target
Free sticky blocks

Construction Tasks

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



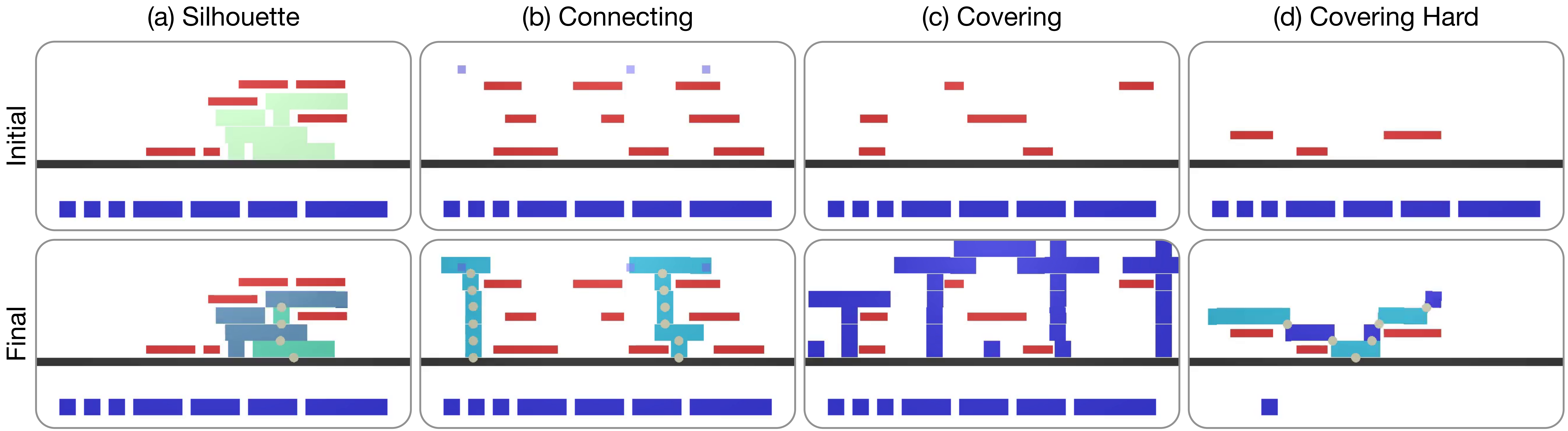
+1 per target
-0.5 per sticky block

+1 per target
Free sticky blocks

Length covered
-2 per sticky block

Construction Tasks

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



(a) Silhouette

(b) Connecting

(c) Covering

(d) Covering Hard

+1 per target
-0.5 per sticky block

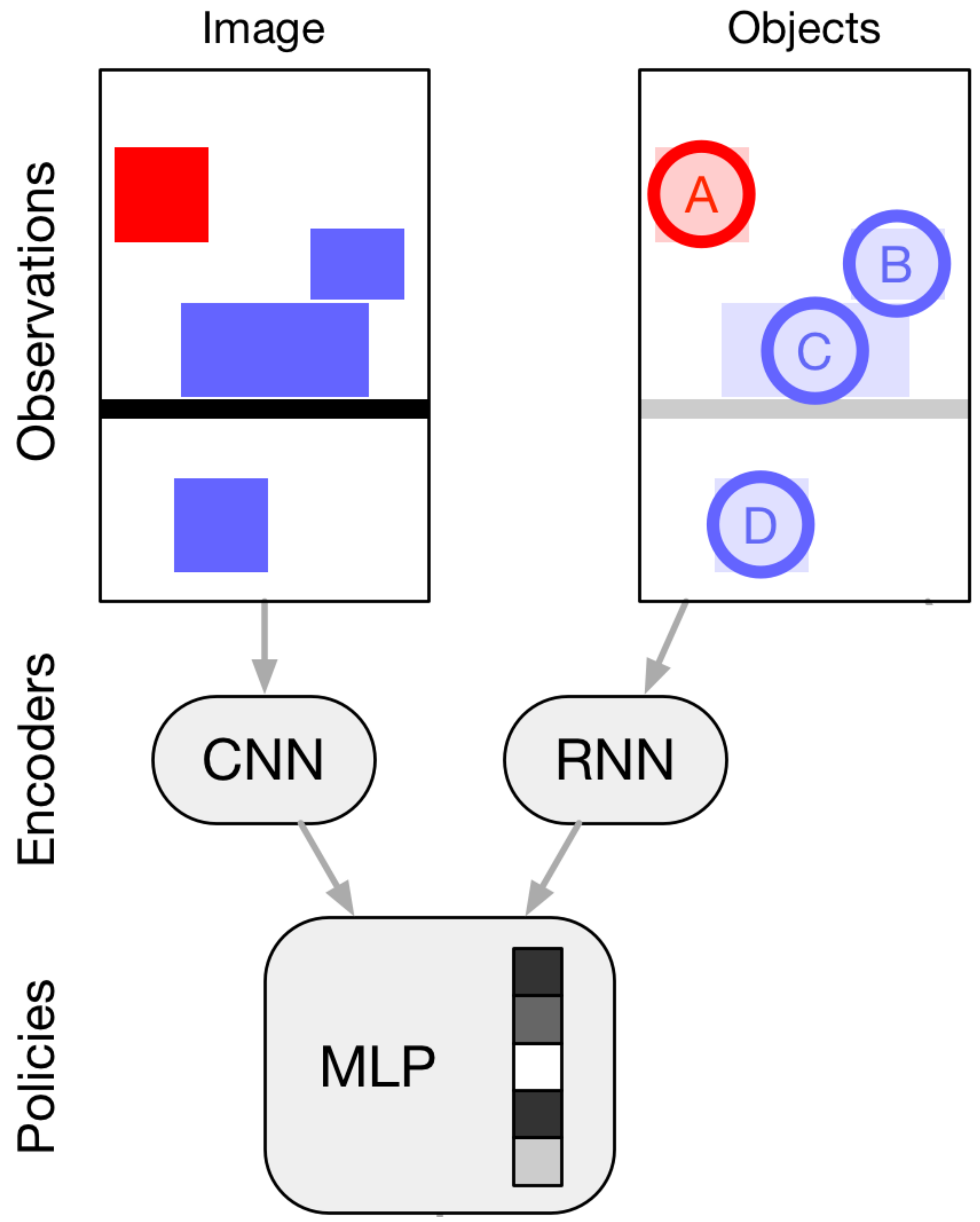
+1 per target
Free sticky blocks

Length covered
-2 per sticky block

Length covered
-0.5 per sticky block

Absolute and Relative Actions

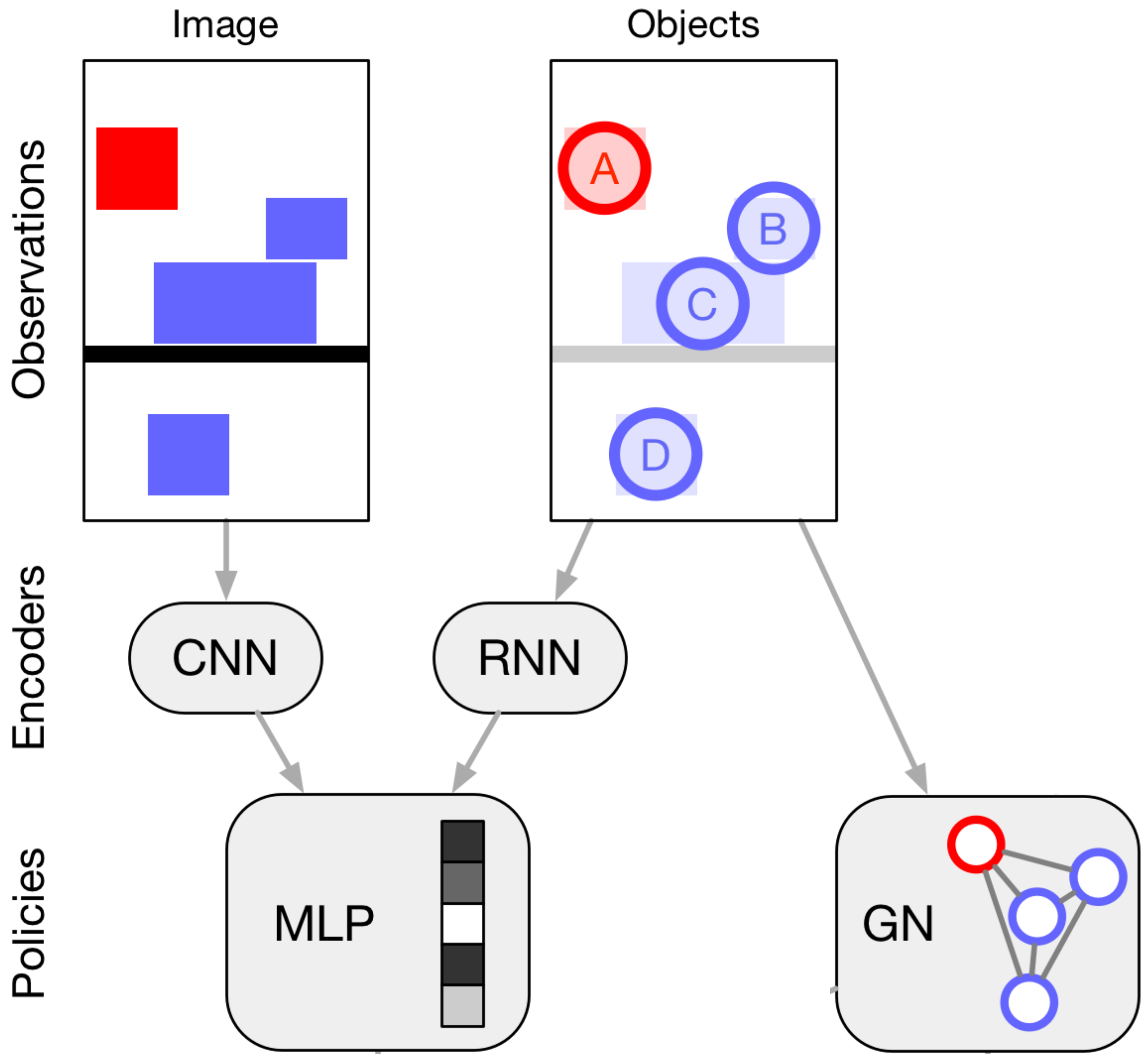
Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



“Place block D at (x, y)”

Absolute and Relative Actions

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

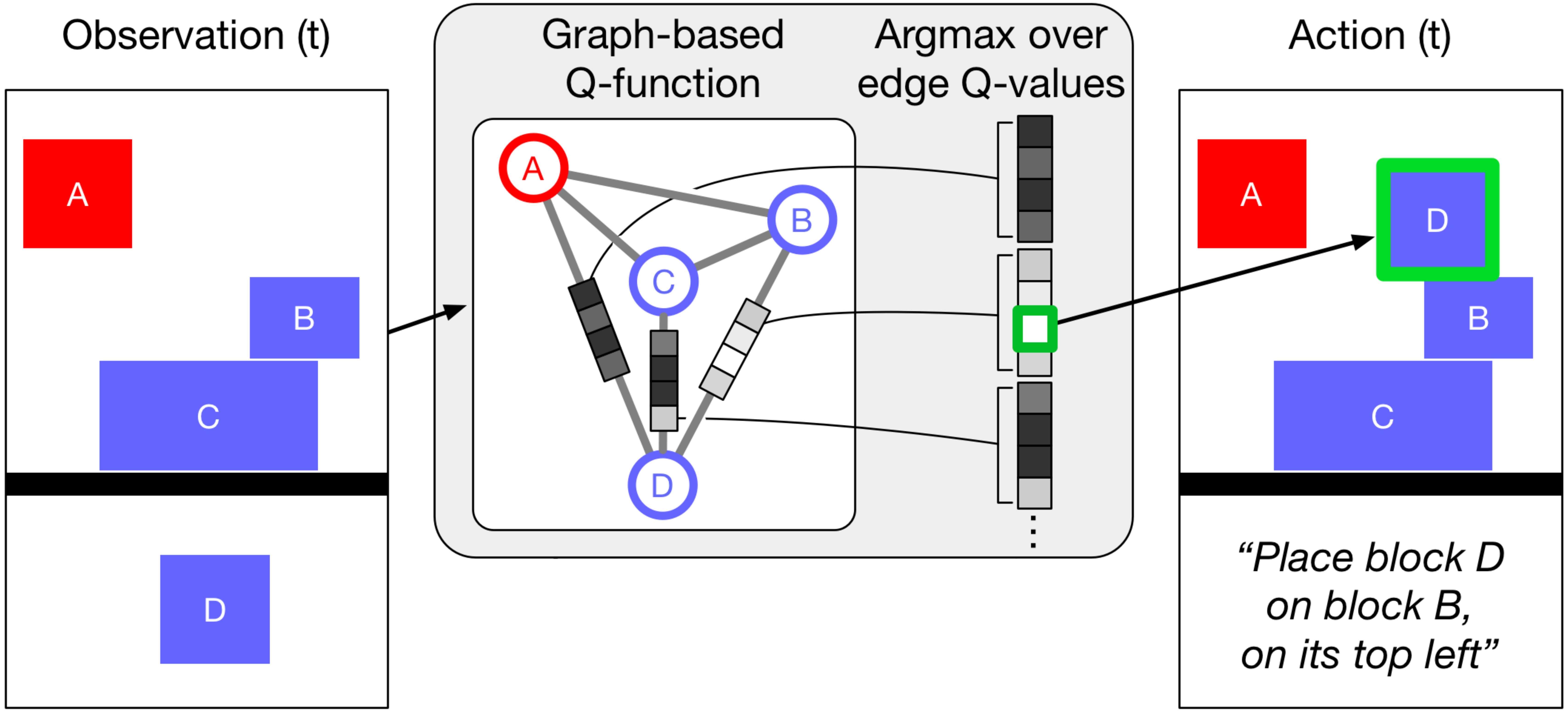


"Place block D at (x, y)"

"Place block D on block B"

Graph Network Agent

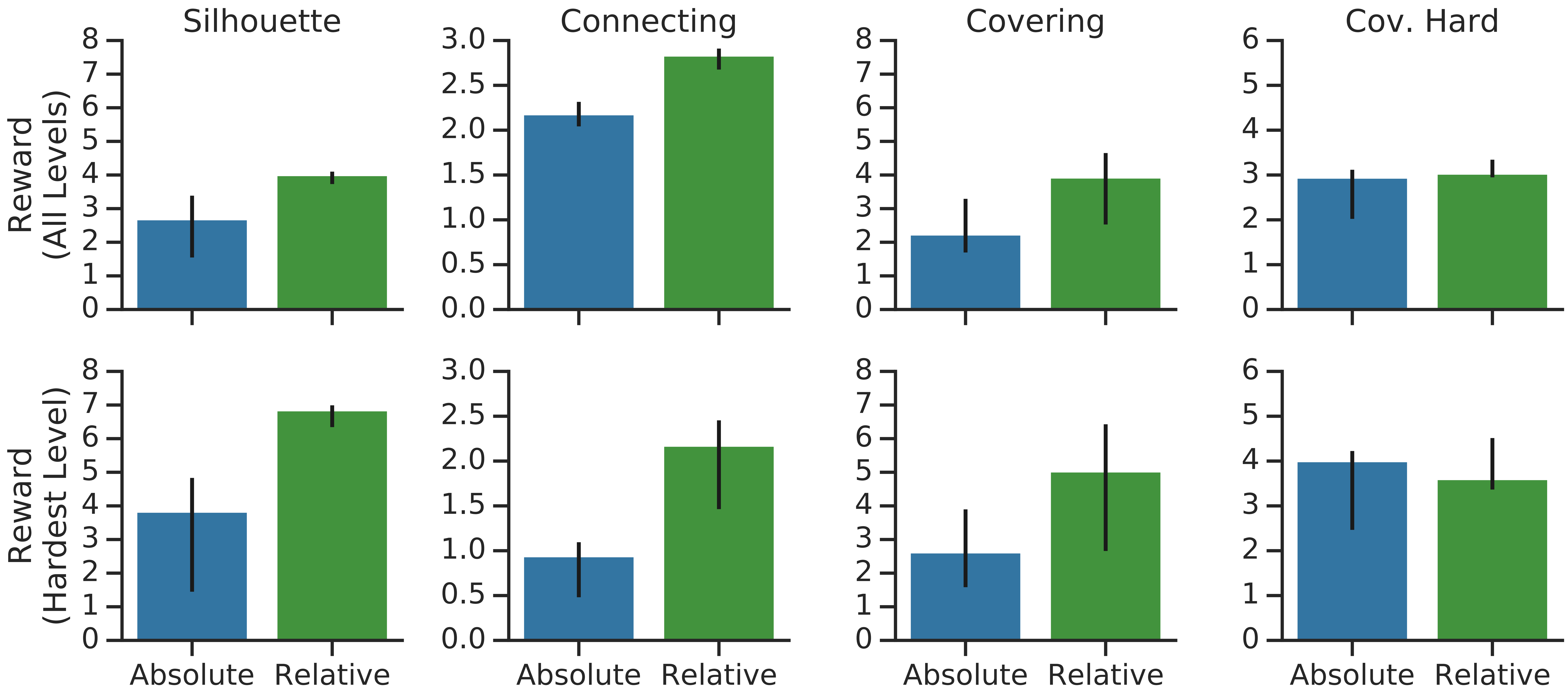
Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



"Place block D on block B, on its top left"

Overall Results

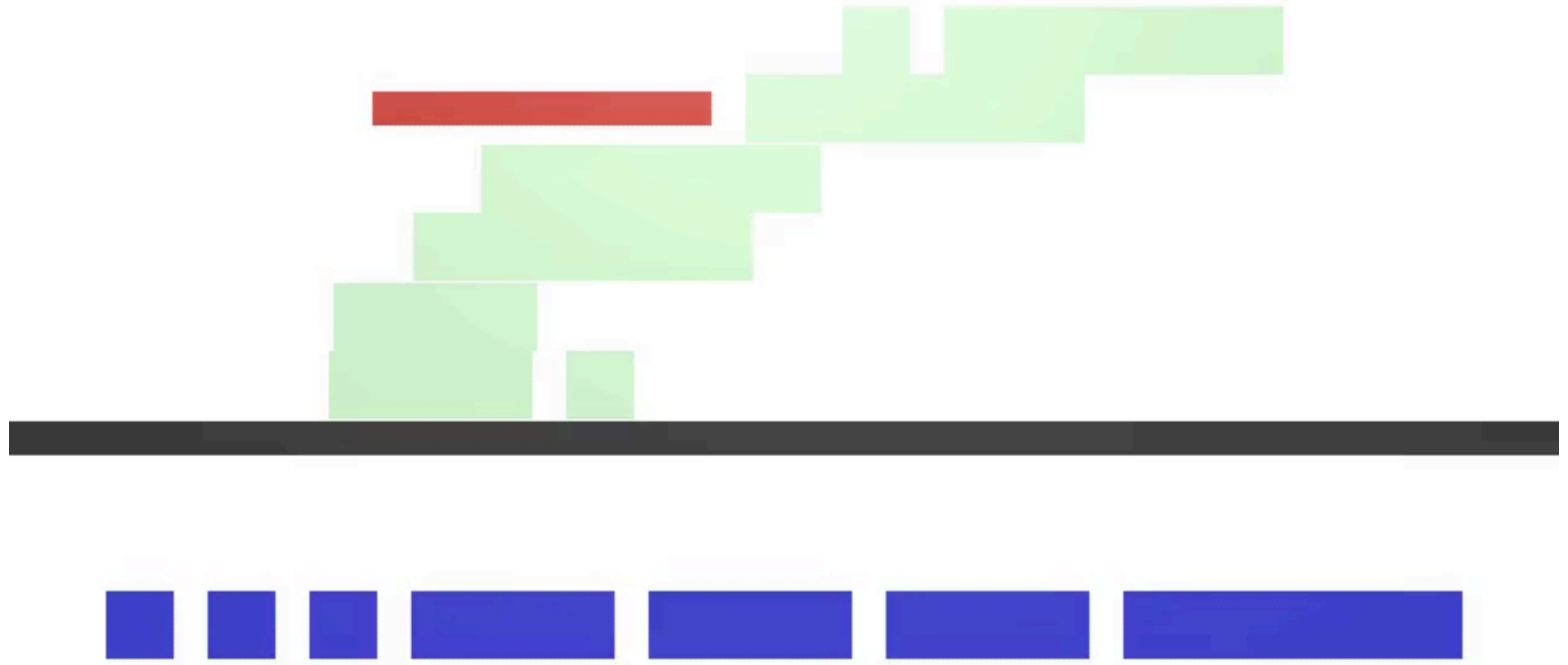
Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



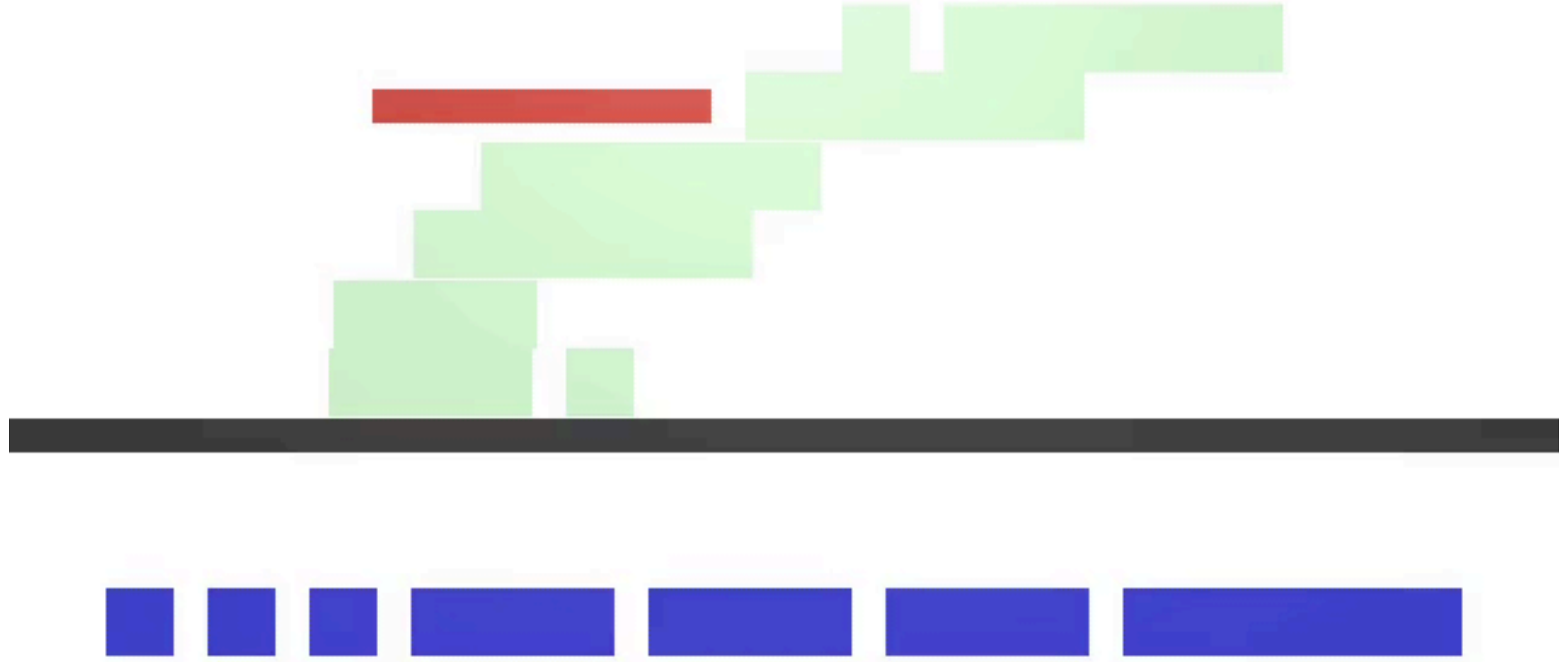
Silhouette

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 4.83)



Object-Centric Actions
(Average reward: 6.99)

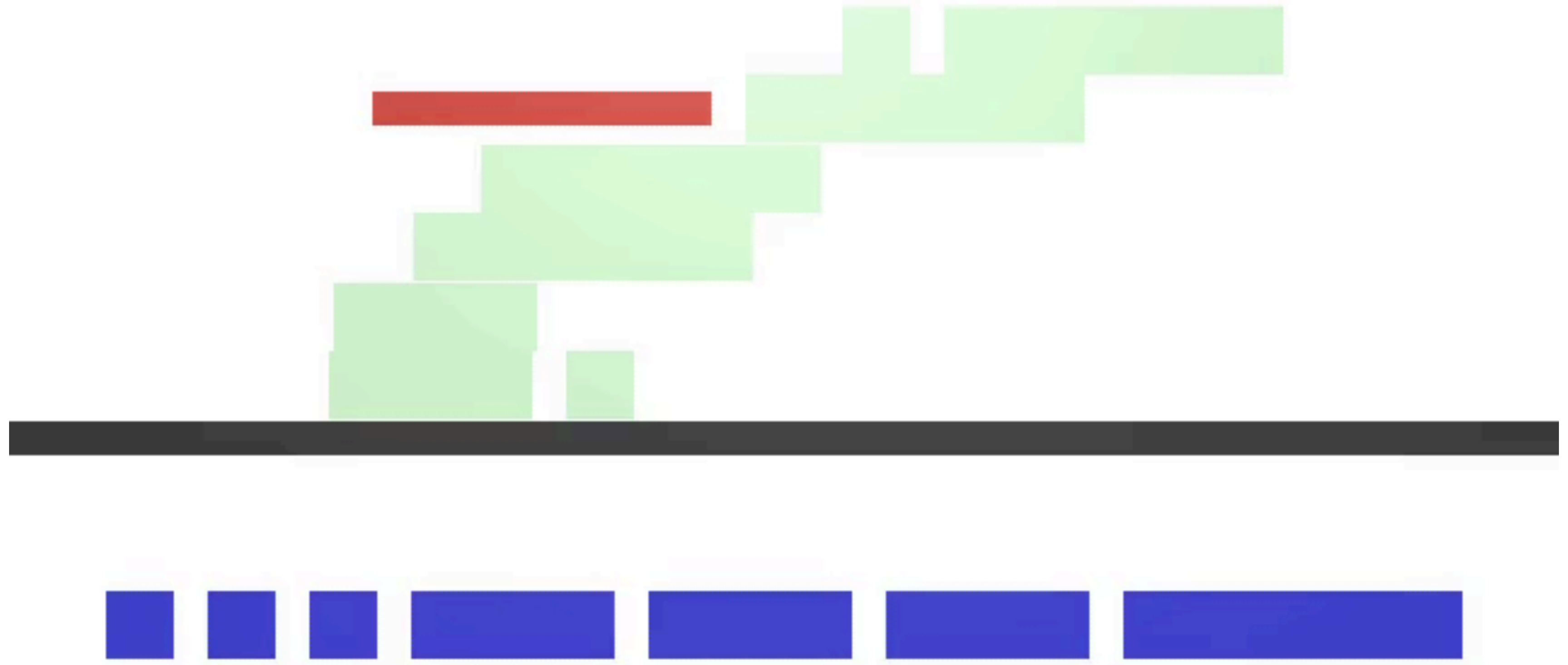


Reward: +1 per target, -0.5 per sticky block

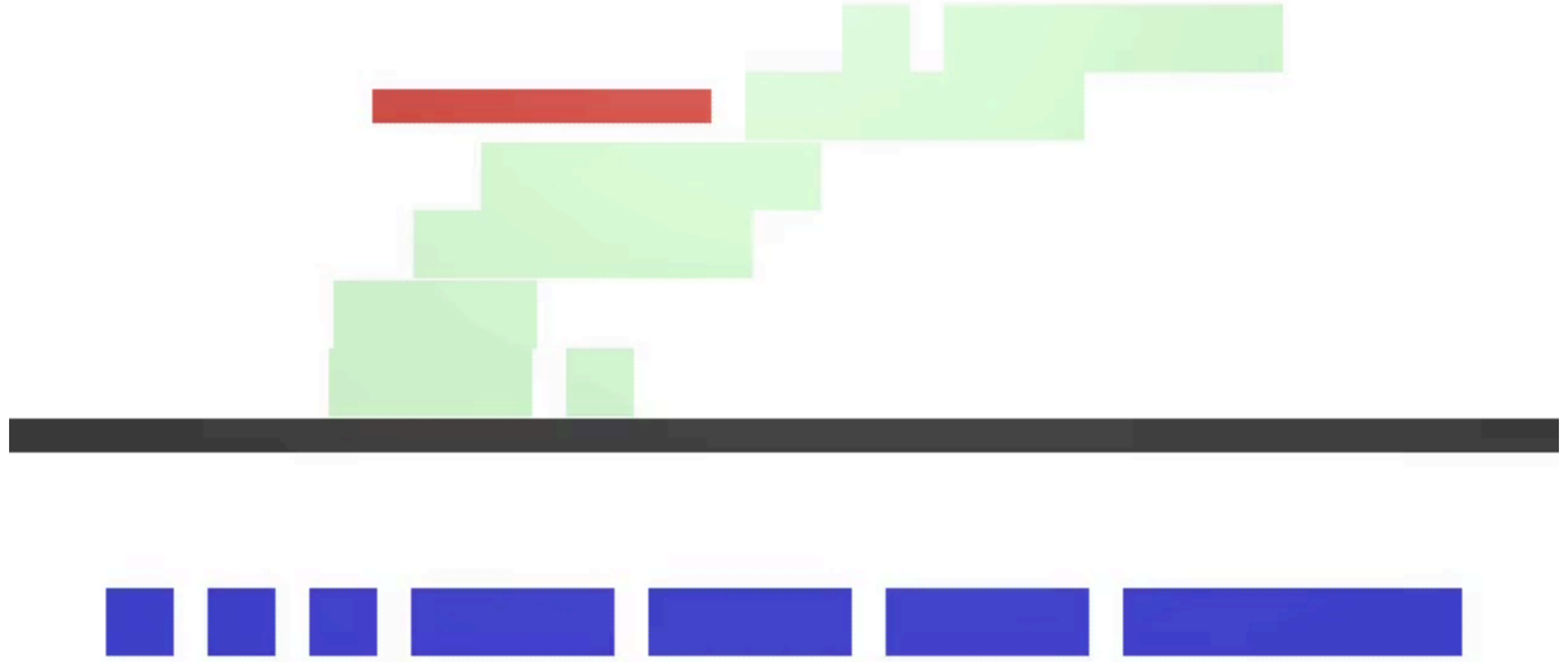
Silhouette

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 4.83)



Object-Centric Actions
(Average reward: 6.99)

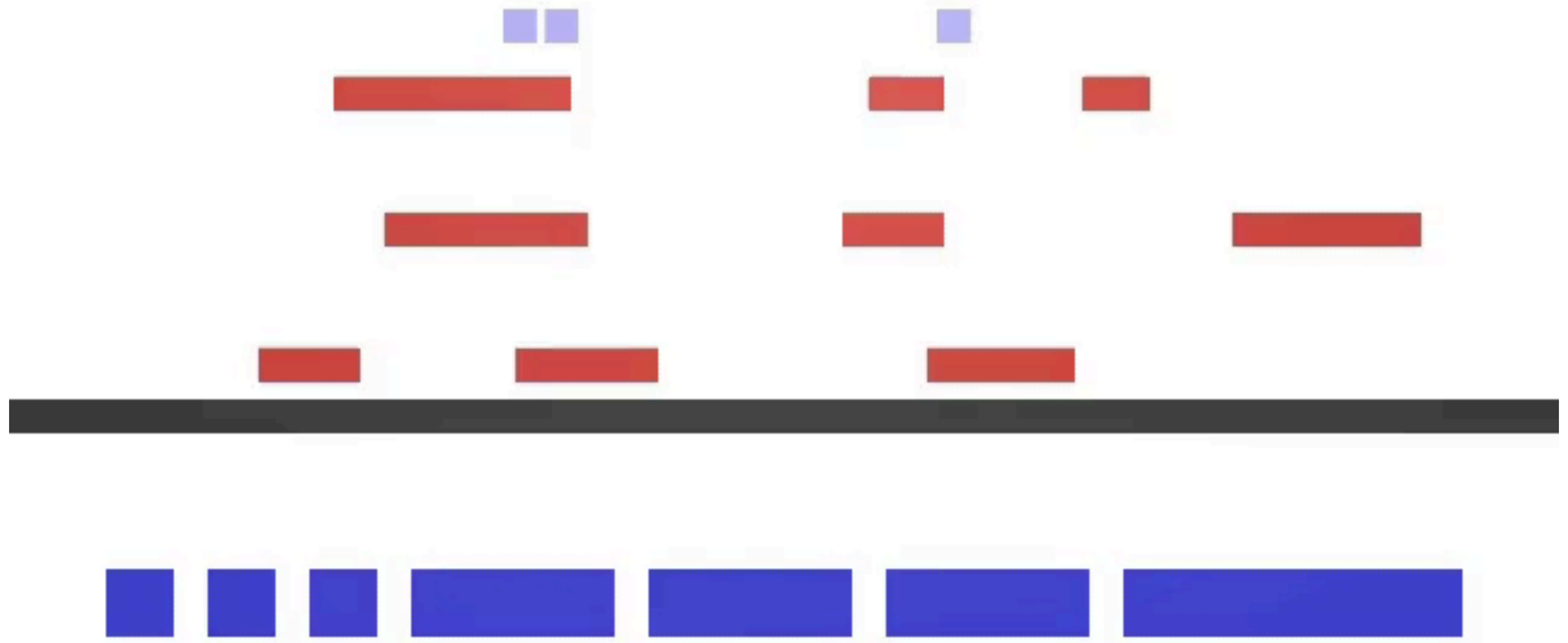


Reward: +1 per target, -0.5 per sticky block

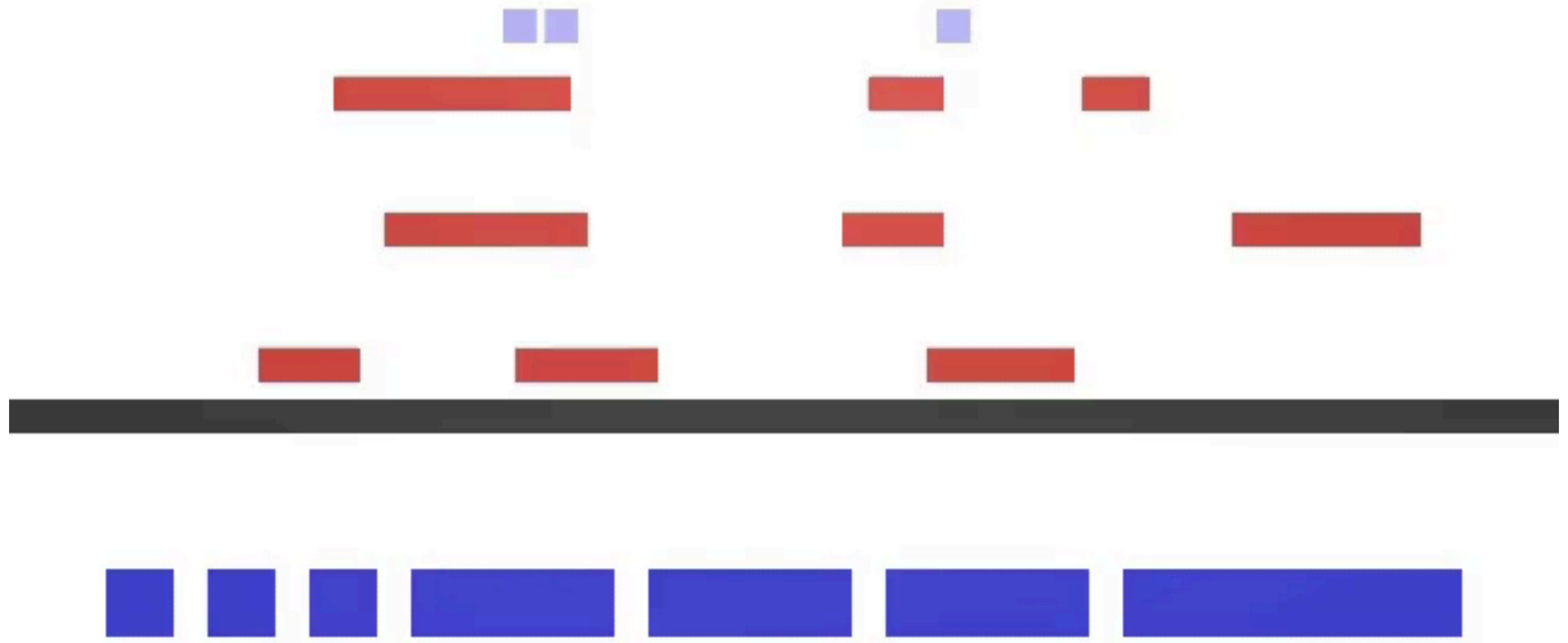
Connecting

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 1.09)



Object-Centric Actions
(Average reward: 2.45)

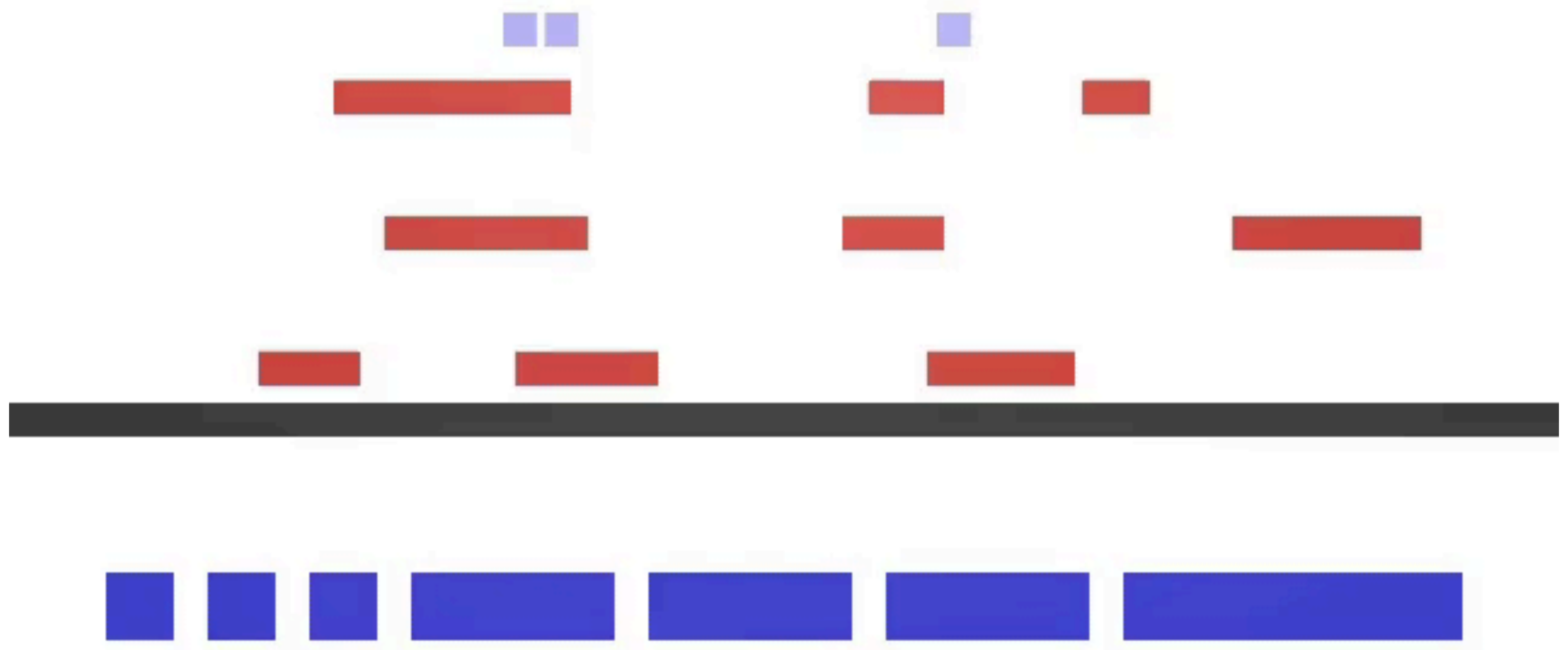


Reward: +1 per target, free sticky blocks

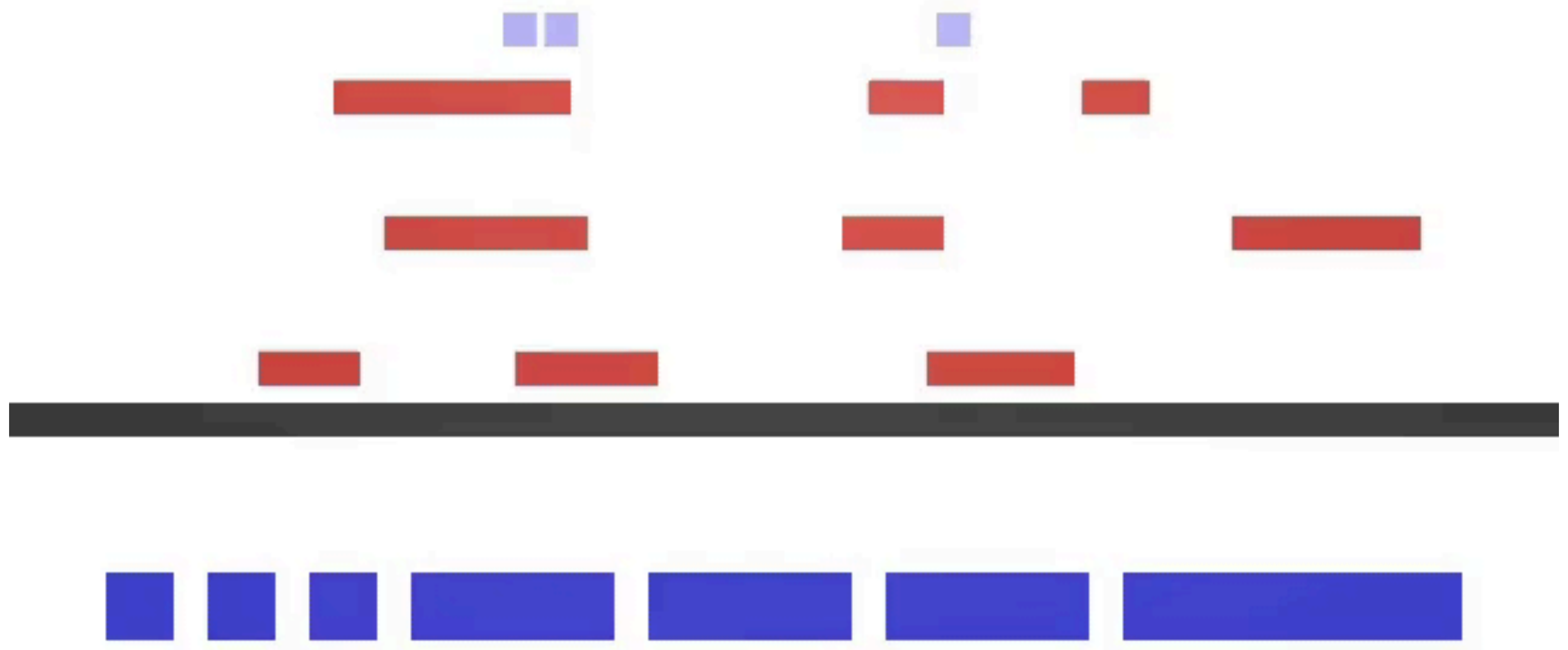
Connecting

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 1.09)



Object-Centric Actions
(Average reward: 2.45)

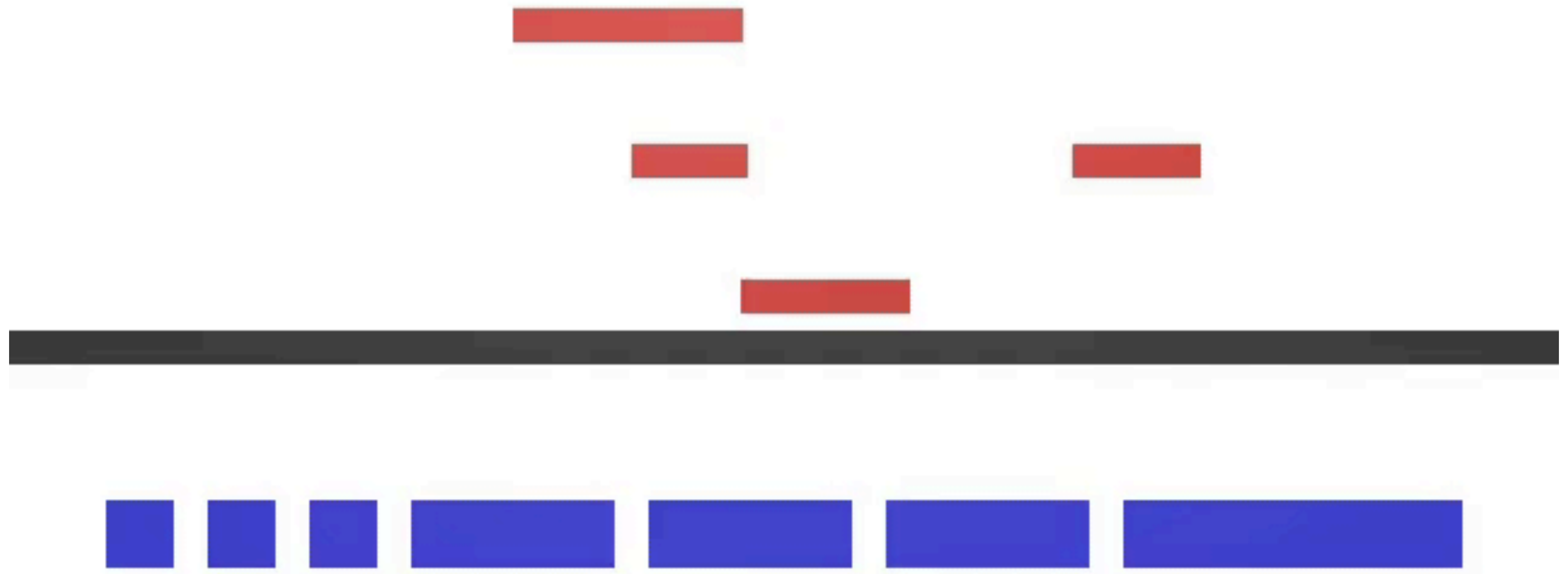


Reward: +1 per target, free sticky blocks

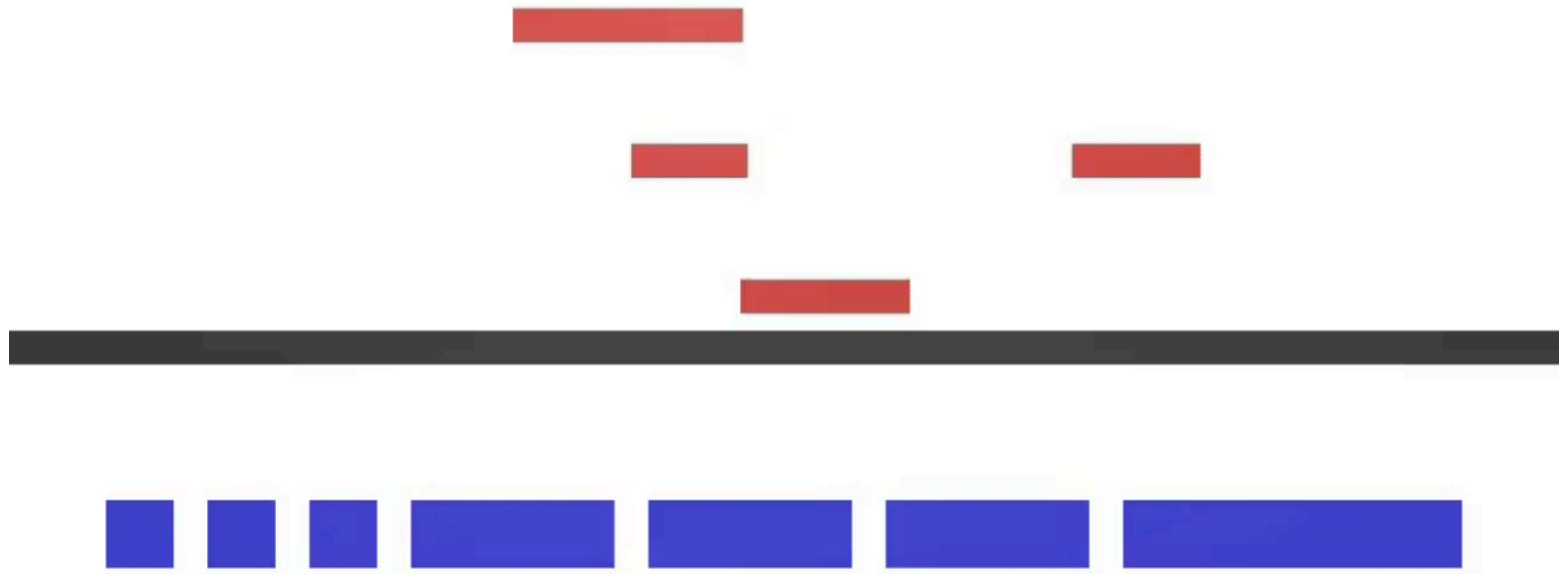
Covering

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 3.90)



Object-Centric Actions
(Average reward: 6.43)

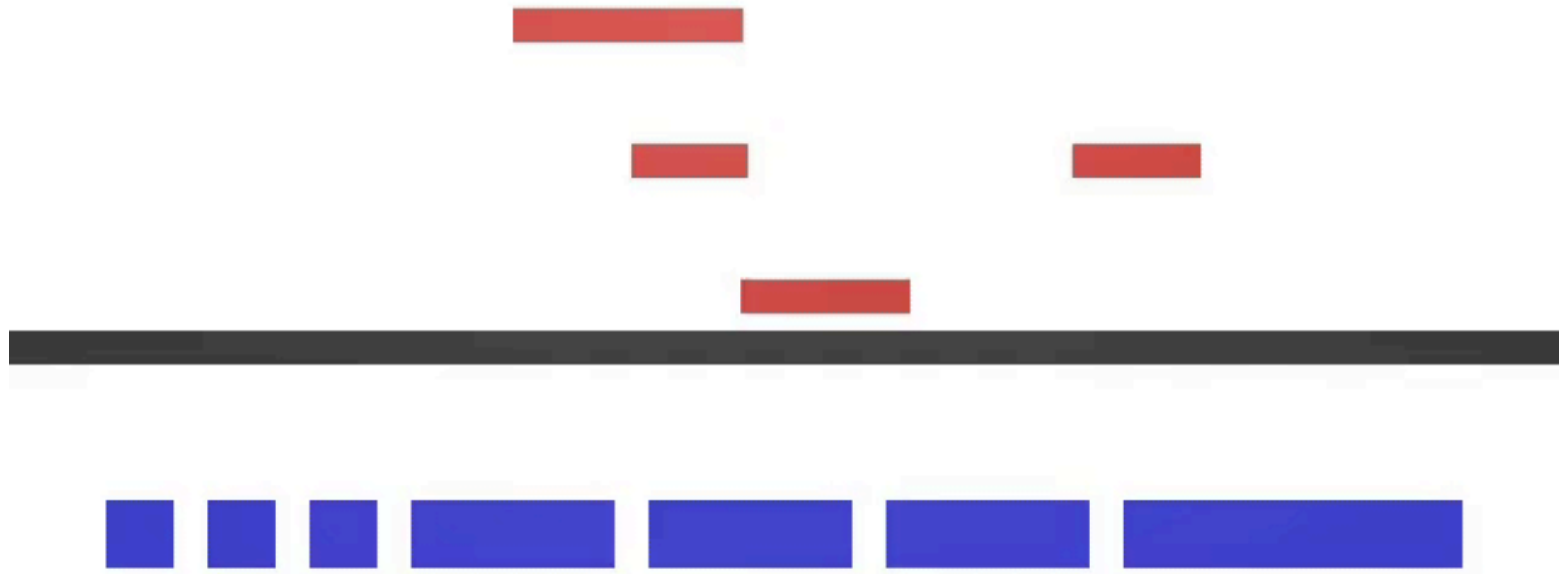


Reward: proportional to length covered, -2 per sticky block

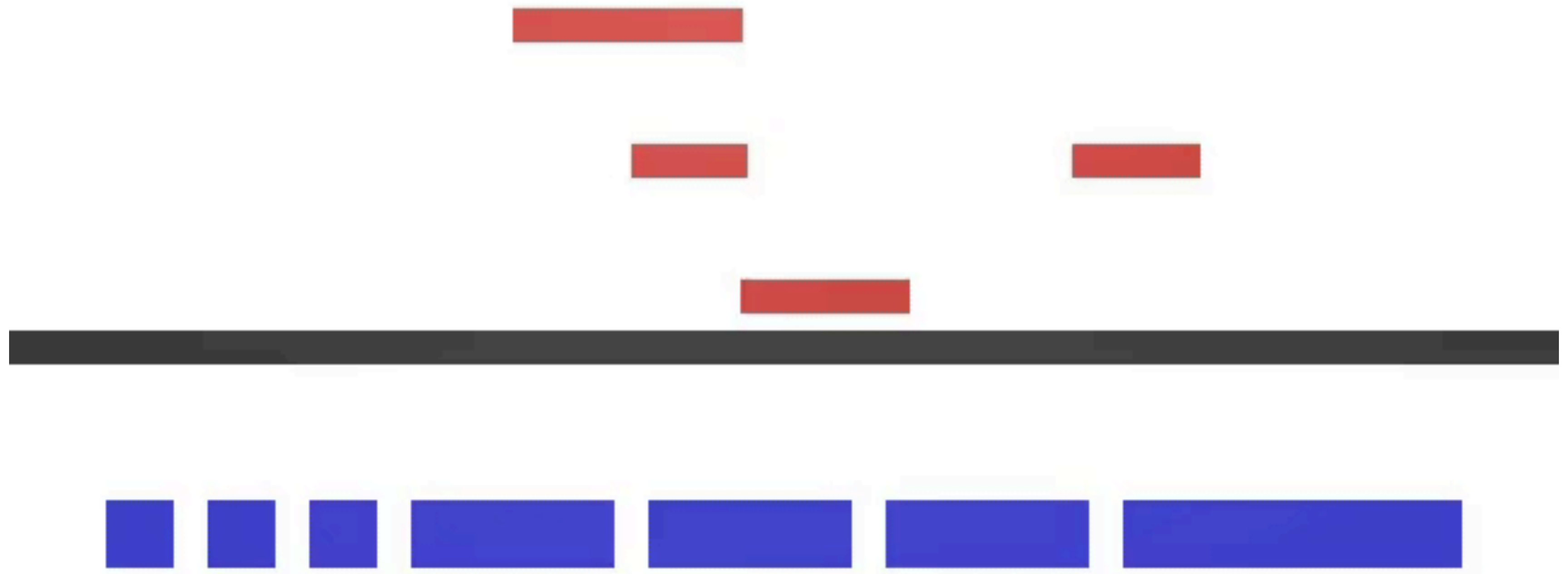
Covering

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 3.90)



Object-Centric Actions
(Average reward: 6.43)

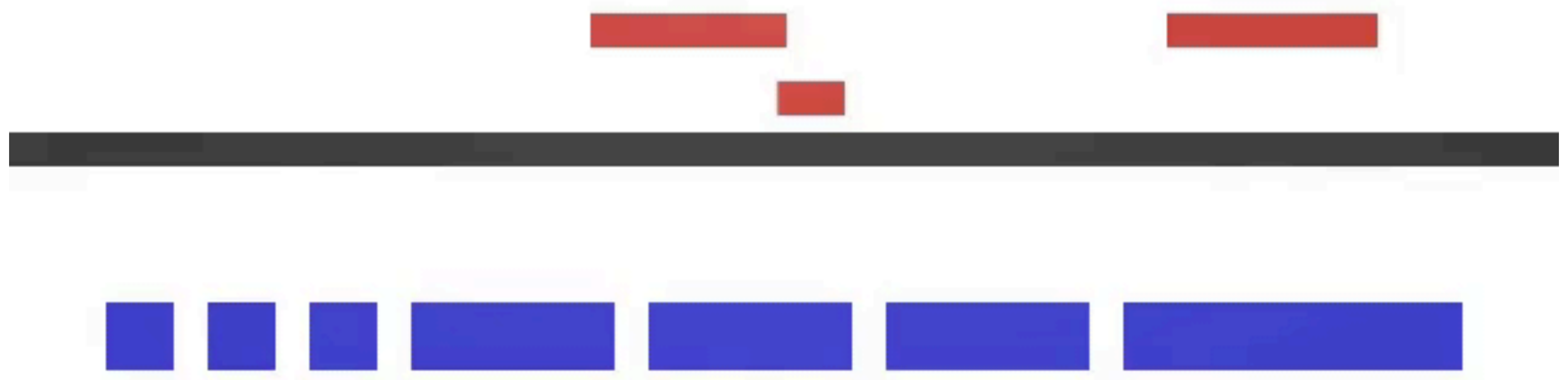


Reward: proportional to length covered, -2 per sticky block

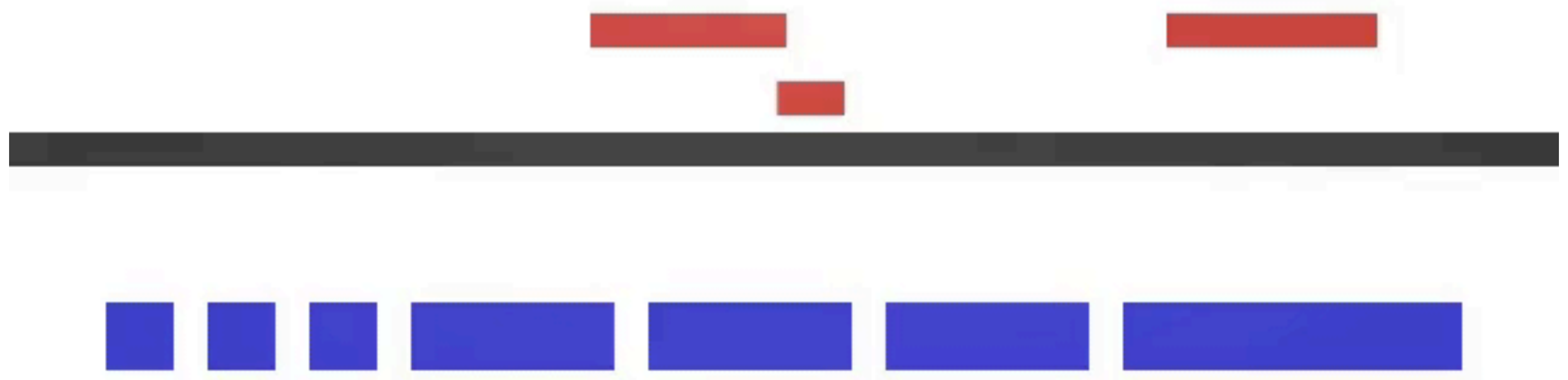
Covering Hard

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 4.22)



Object-Centric Actions
(Average reward: 4.51)

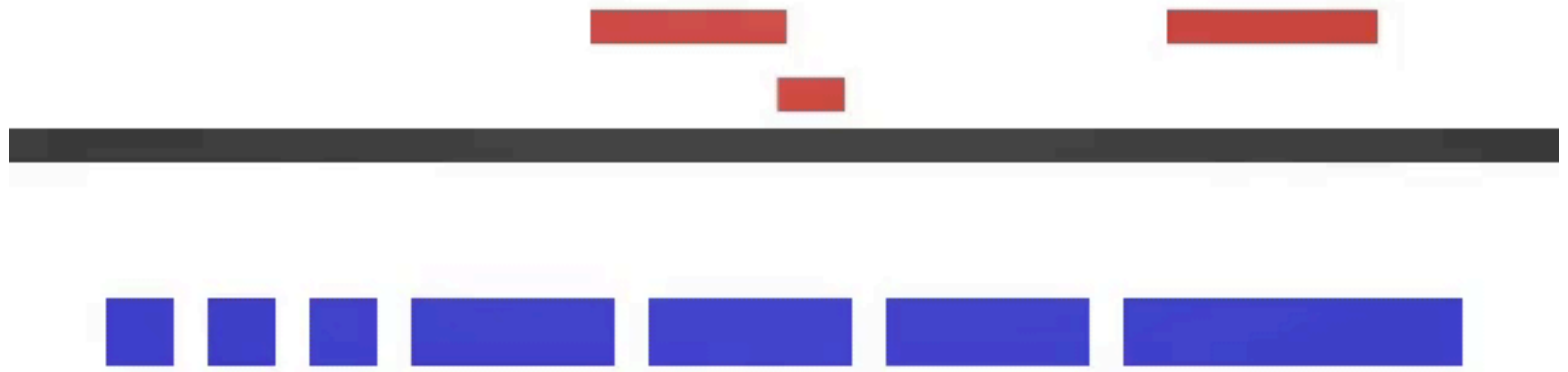


Reward: proportional to length covered, -0.5 per sticky block

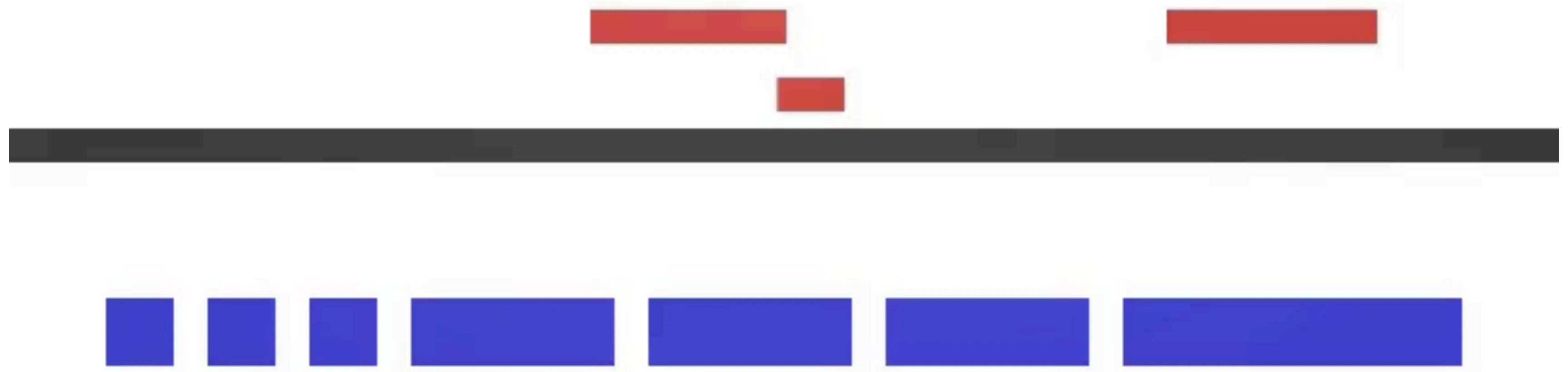
Covering Hard

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Absolute Actions
(Average reward: 4.22)



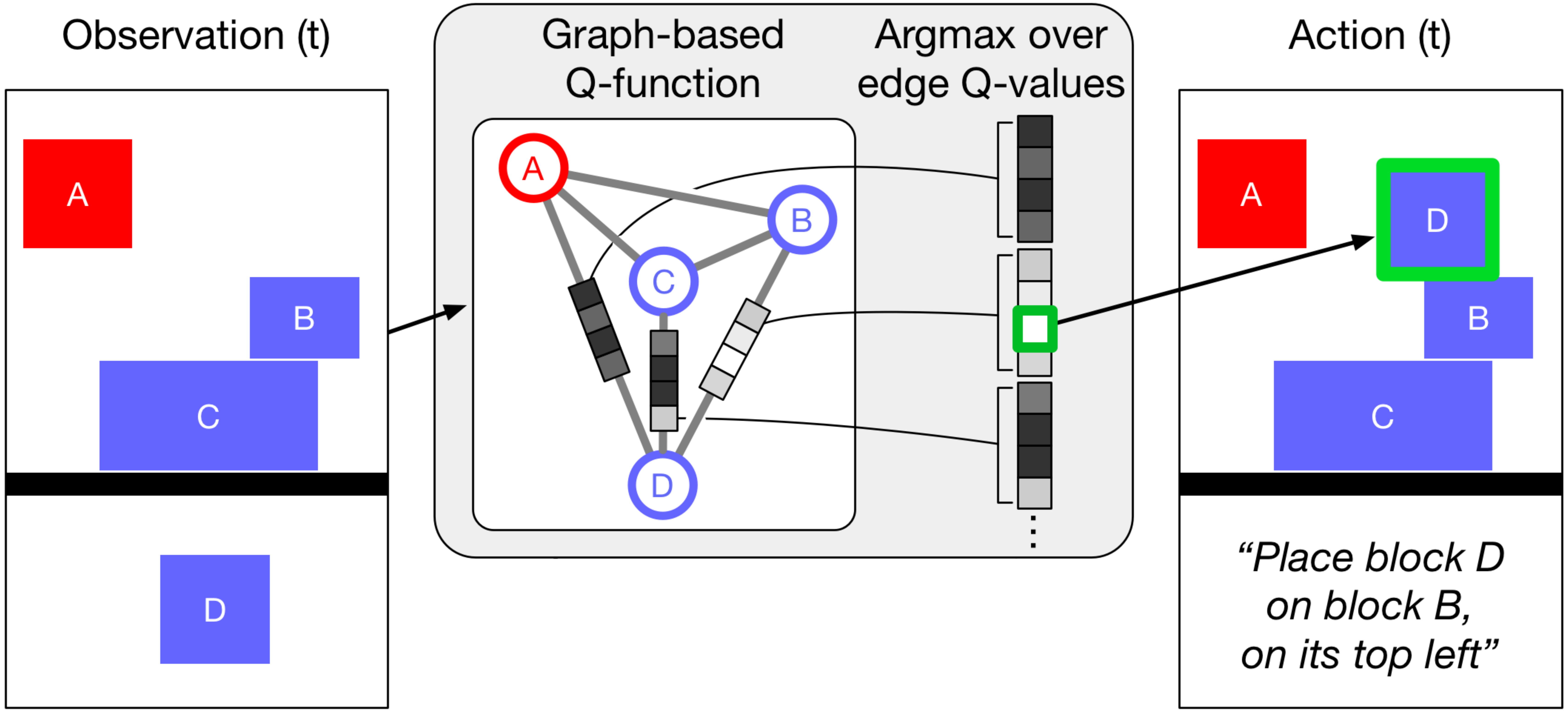
Object-Centric Actions
(Average reward: 4.51)



Reward: proportional to length covered, -0.5 per sticky block

Graph Network Agent

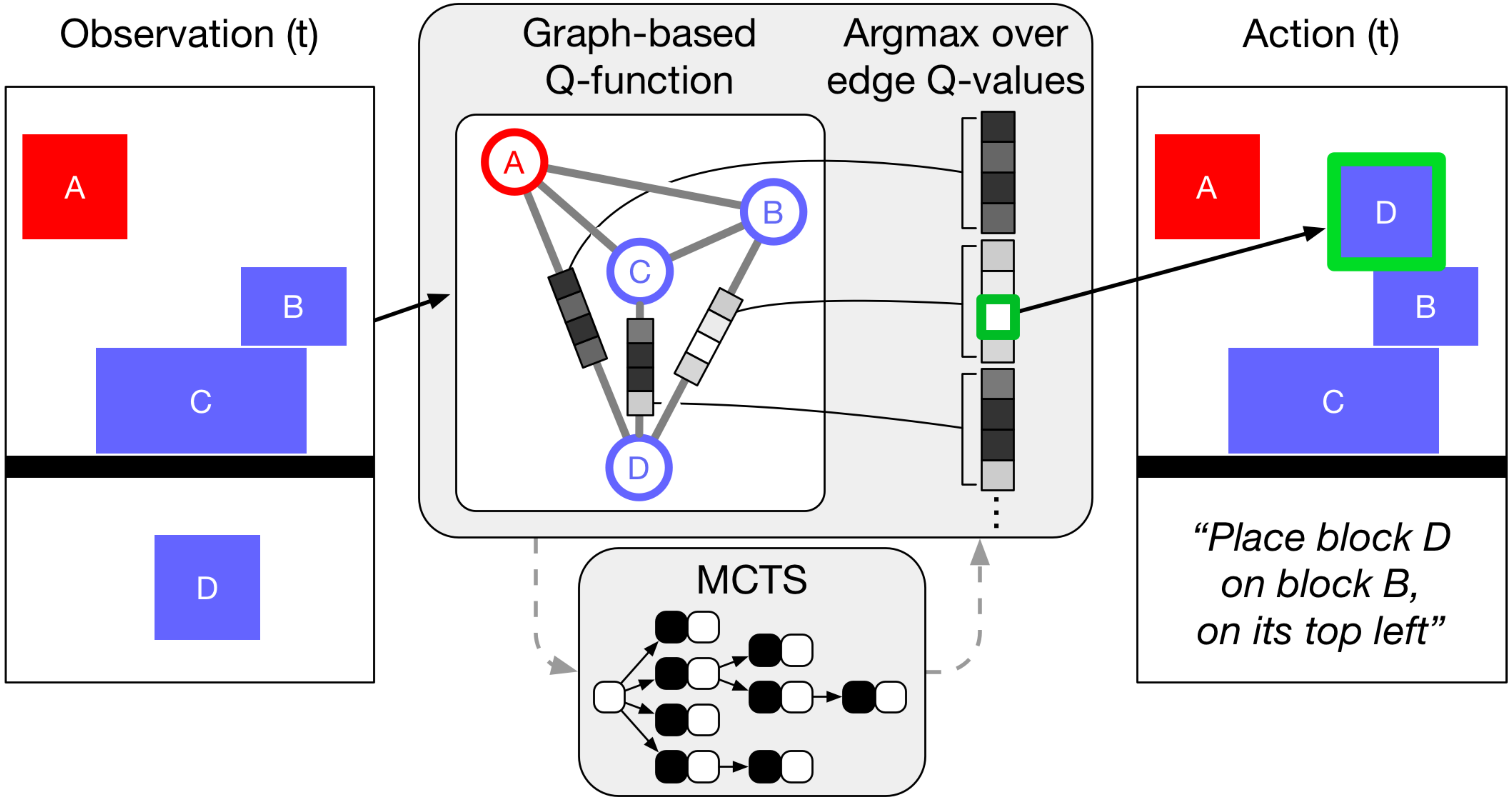
Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



"Place block D on block B, on its top left"

Graph Network Agent

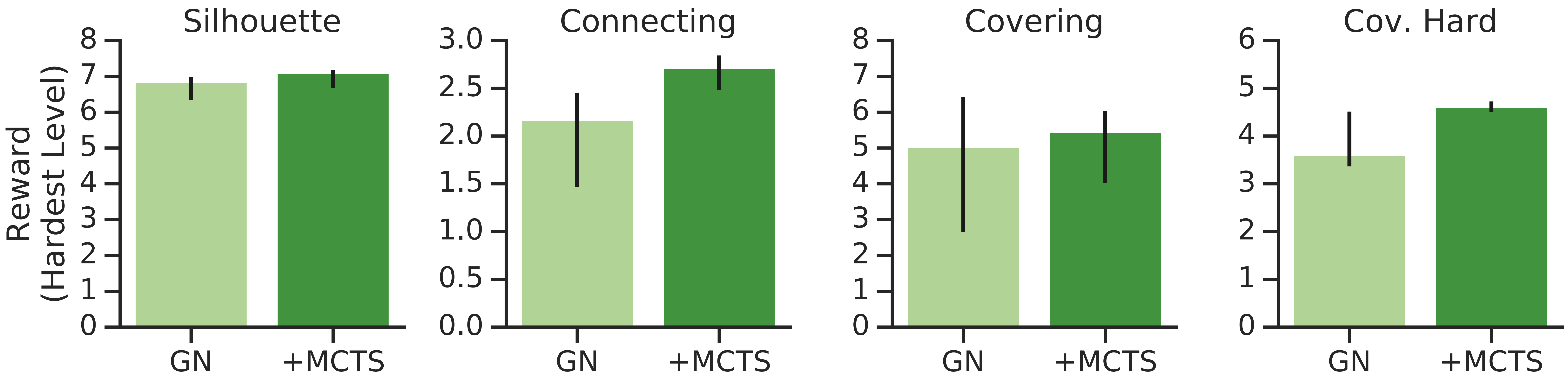
Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



“Place block D on block B, on its top left”

Results: Planning

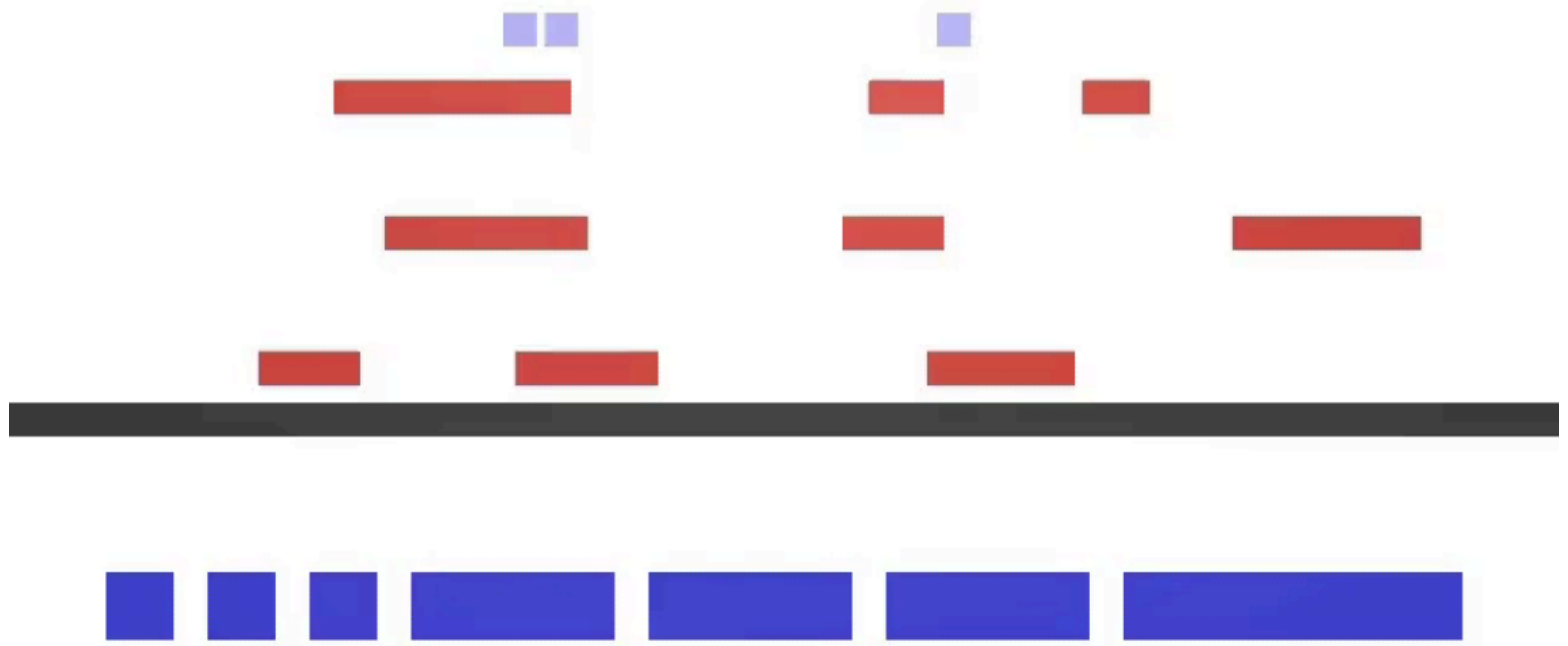
Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



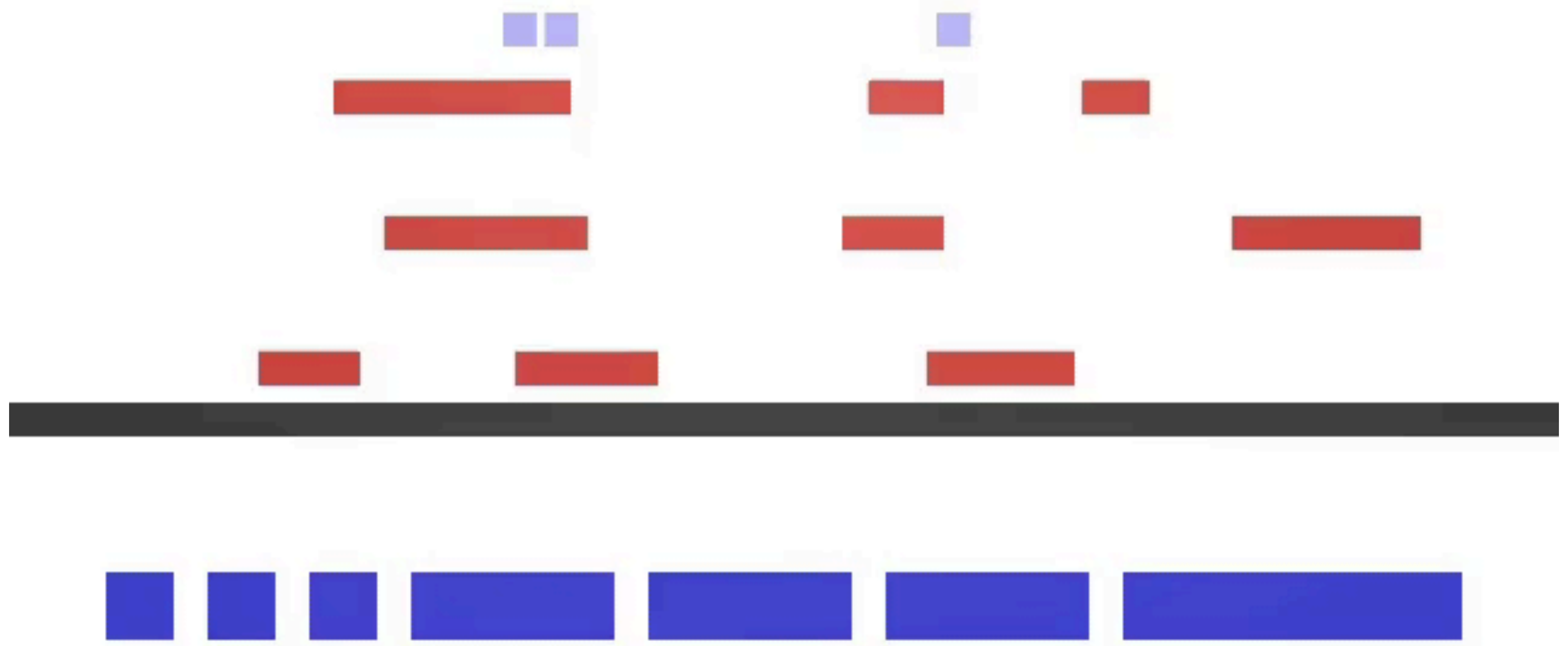
Results: Planning

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Object-Centric Actions
(Average reward: 2.45)



Object-Centric Actions +MCTS
(Average reward: 2.84)

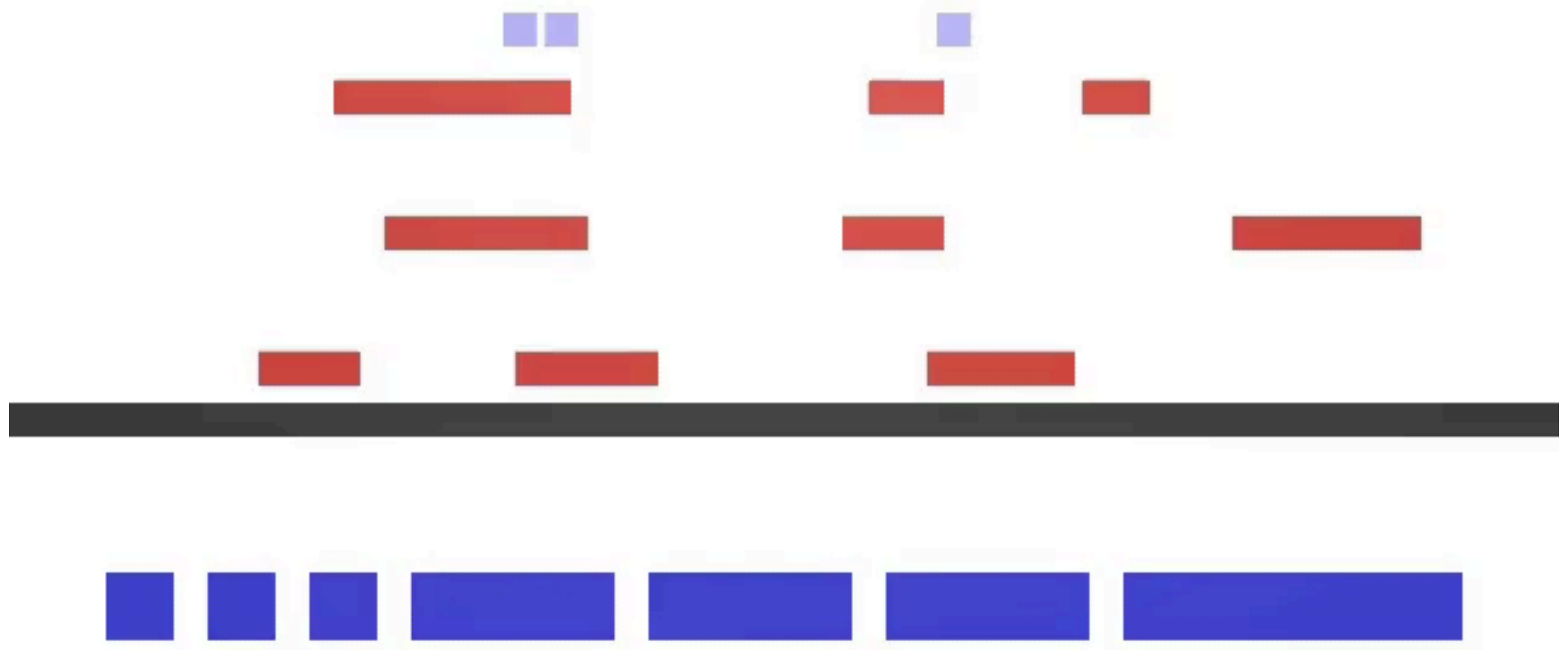


Reward: +1 per target, free sticky blocks

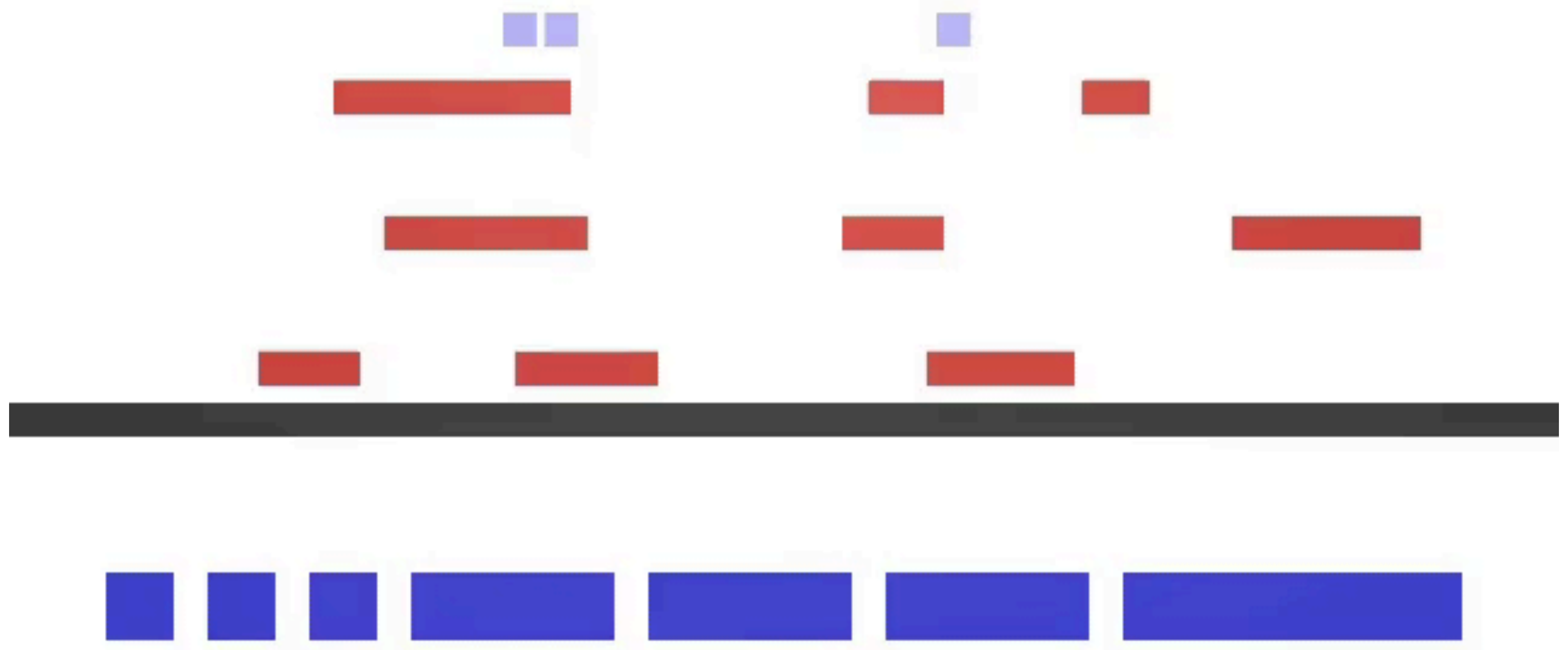
Results: Planning

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

Object-Centric Actions
(Average reward: 2.45)



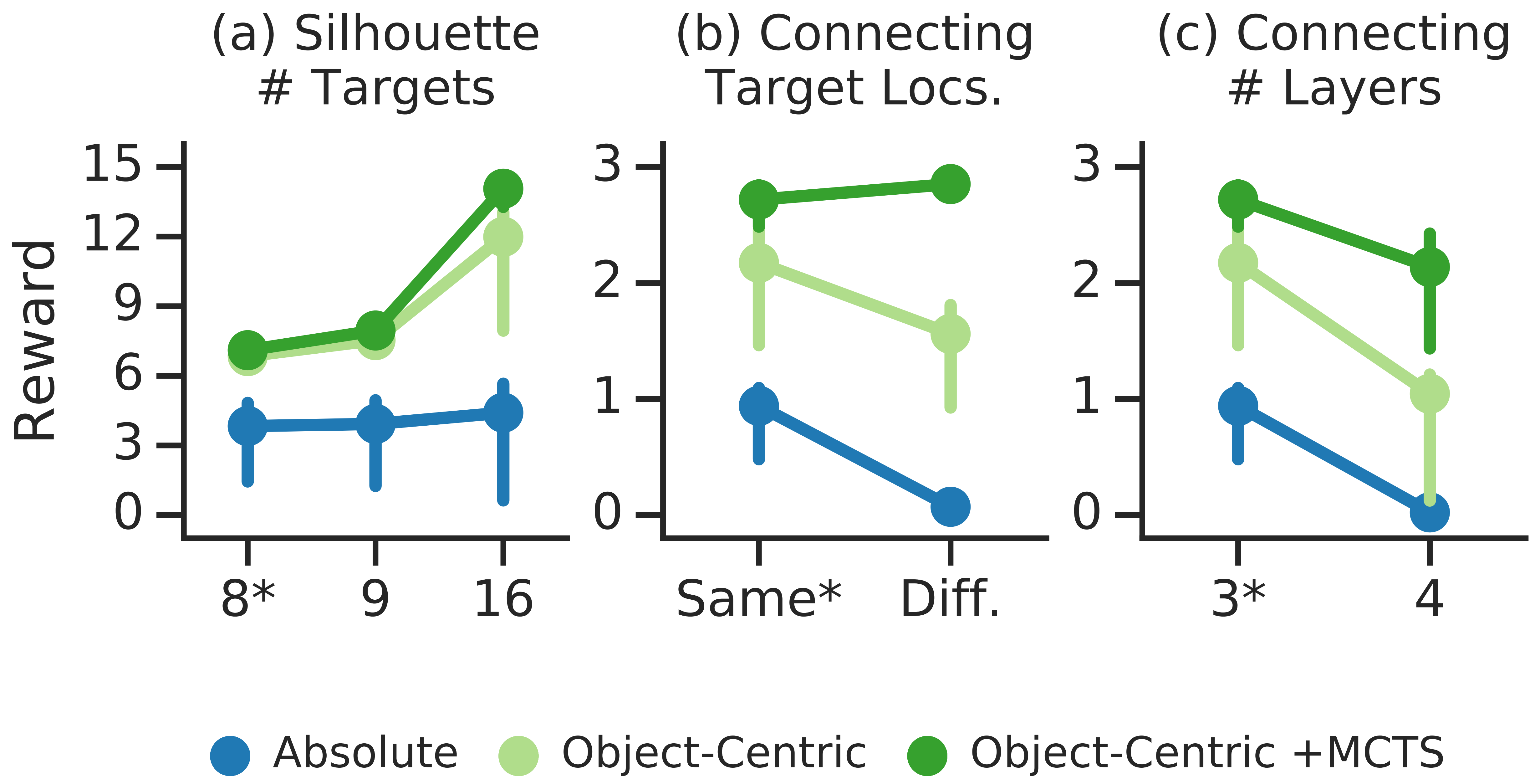
Object-Centric Actions +MCTS
(Average reward: 2.84)



Reward: +1 per target, free sticky blocks

Results: Generalization

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

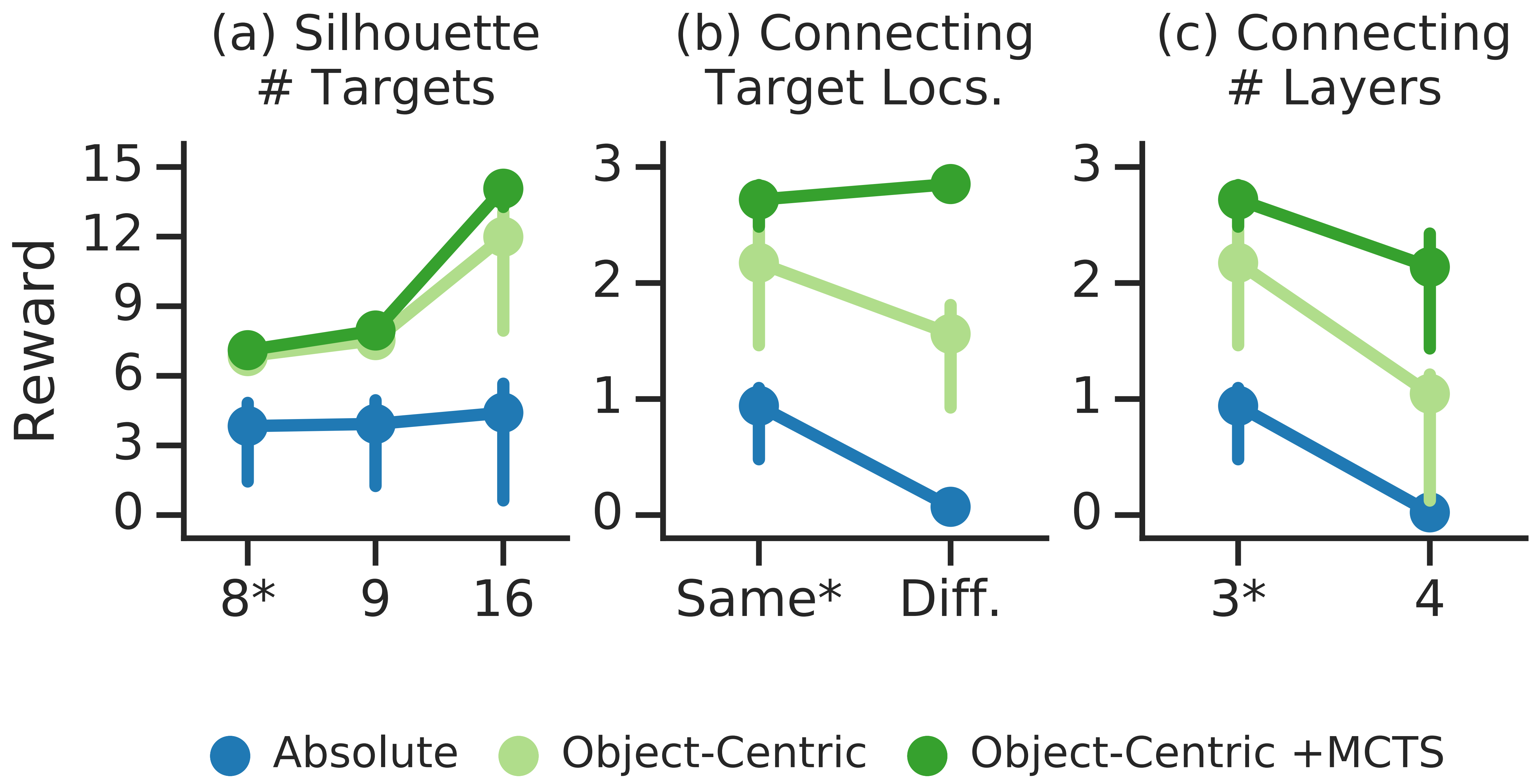


DeepMind

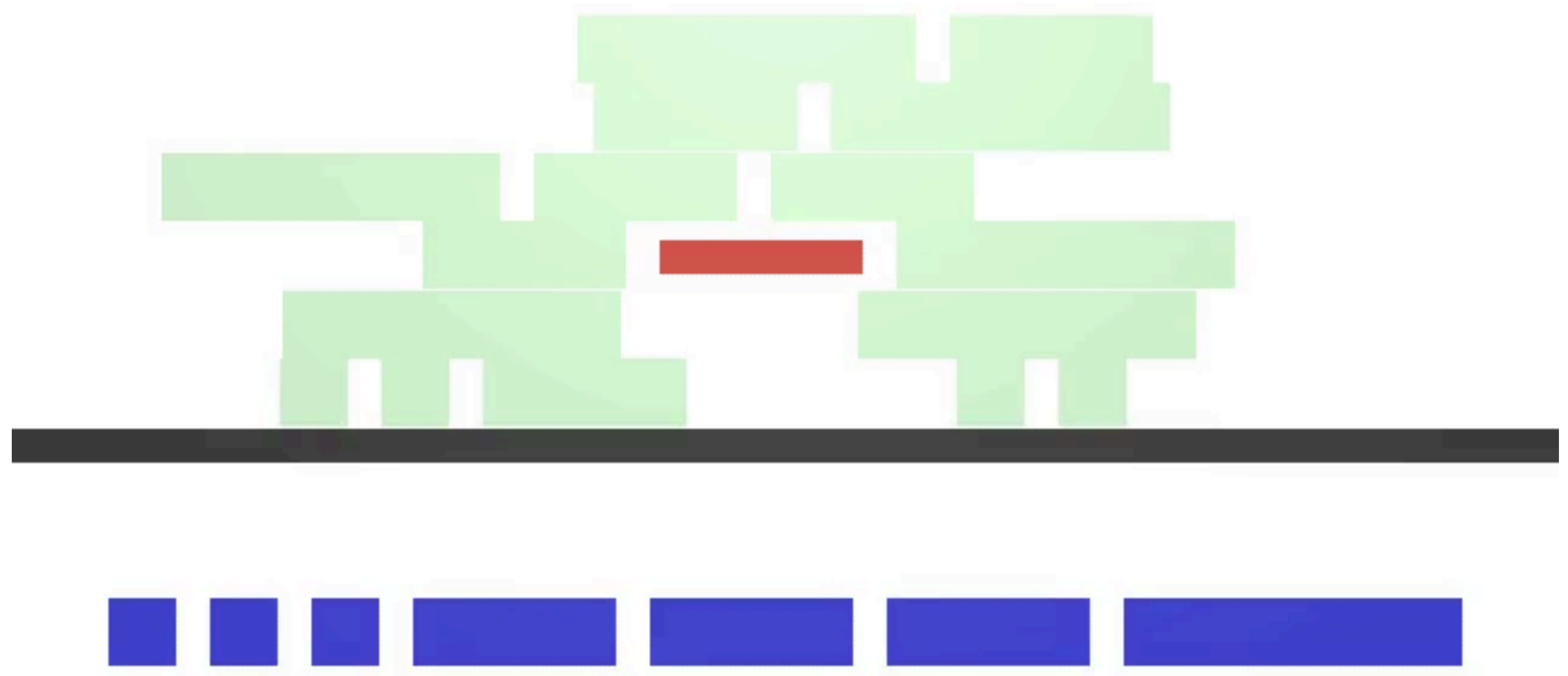
Jessica Hamrick (@jhamrick)

Results: Generalization

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)

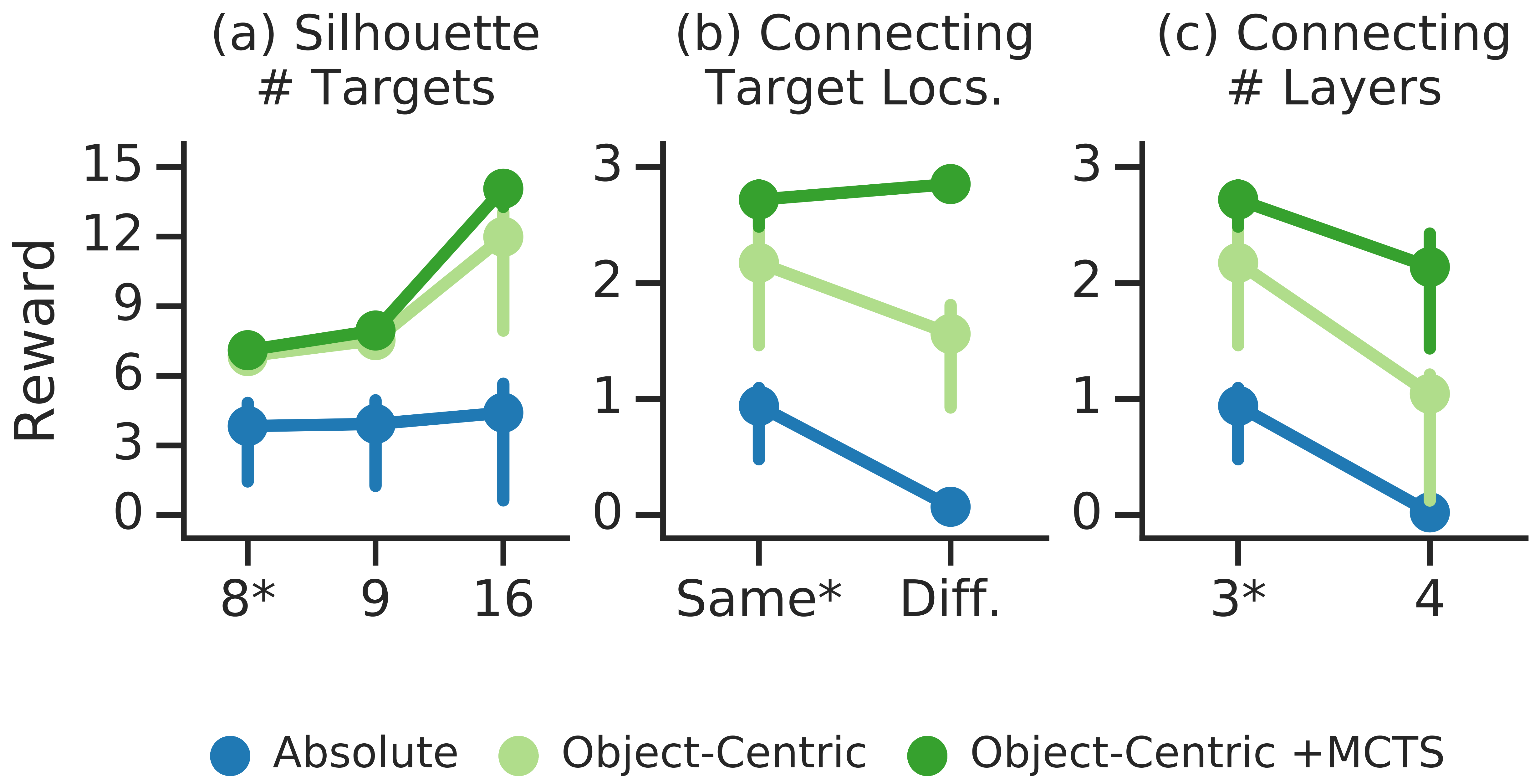


Object-Centric Actions +MCTS
(Average reward: 14.25)

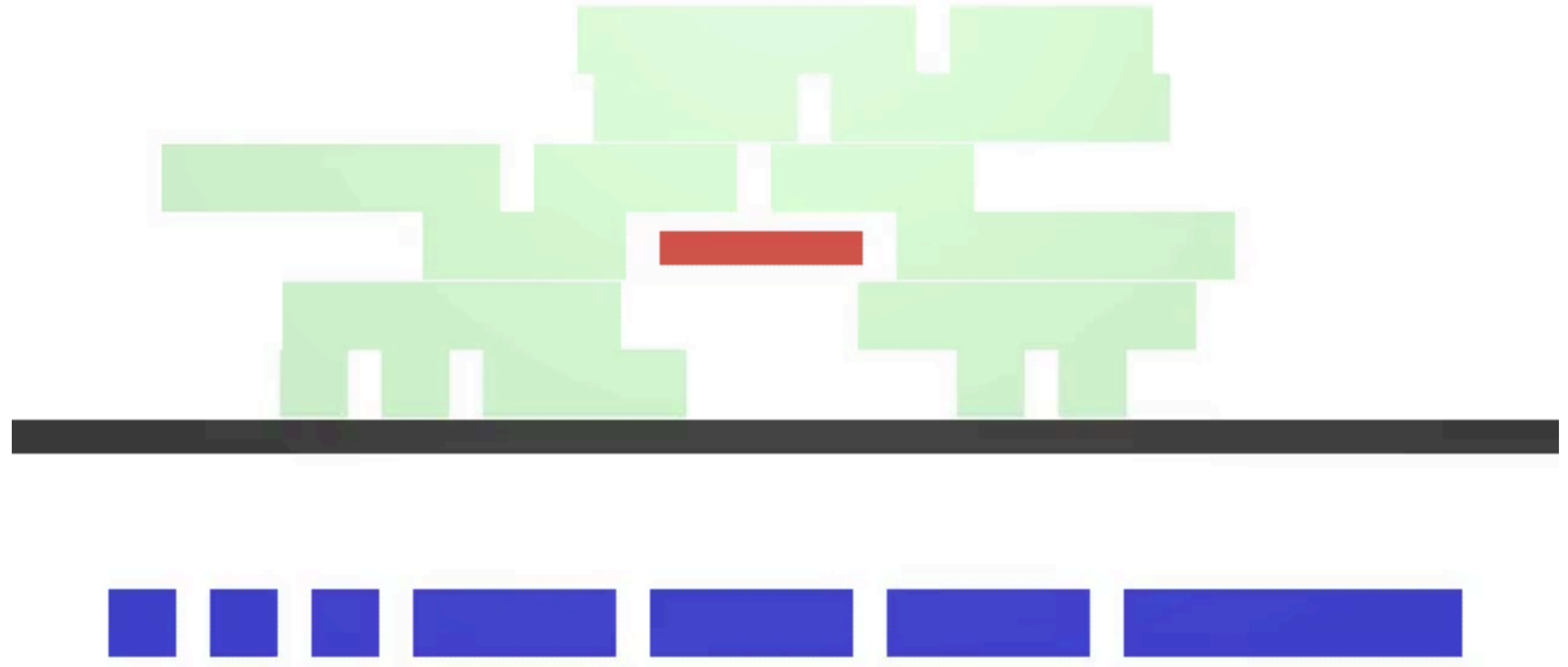


Results: Generalization

Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia & Hamrick (*ICML*, 2019)



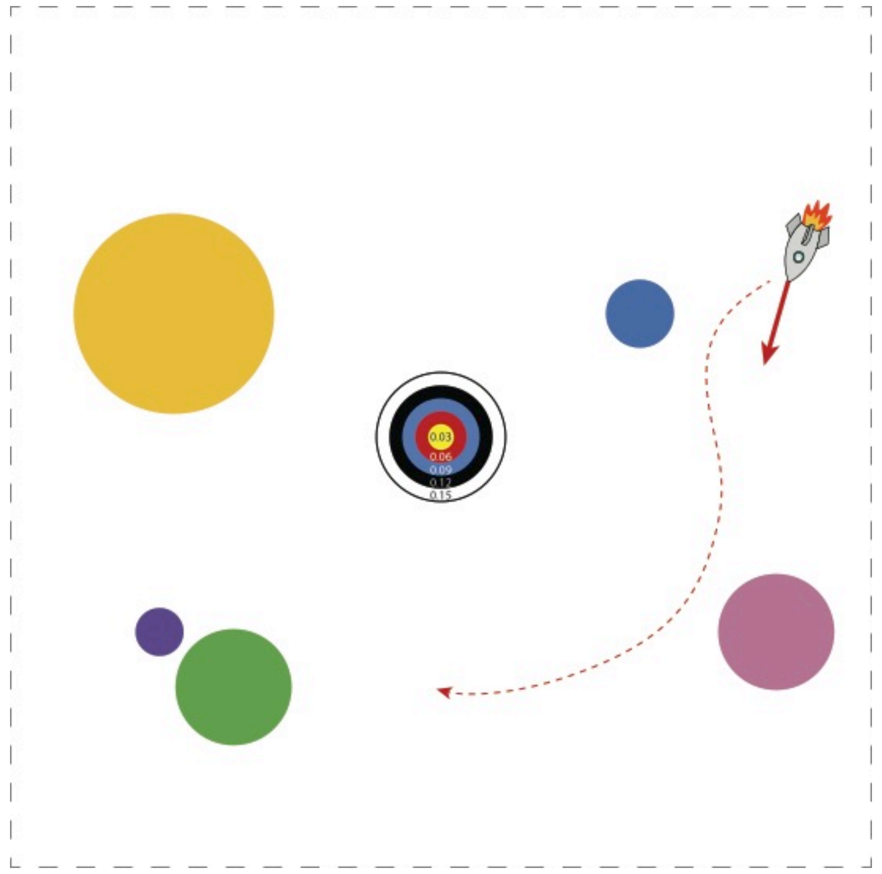
Object-Centric Actions +MCTS
(Average reward: 14.25)



Outline

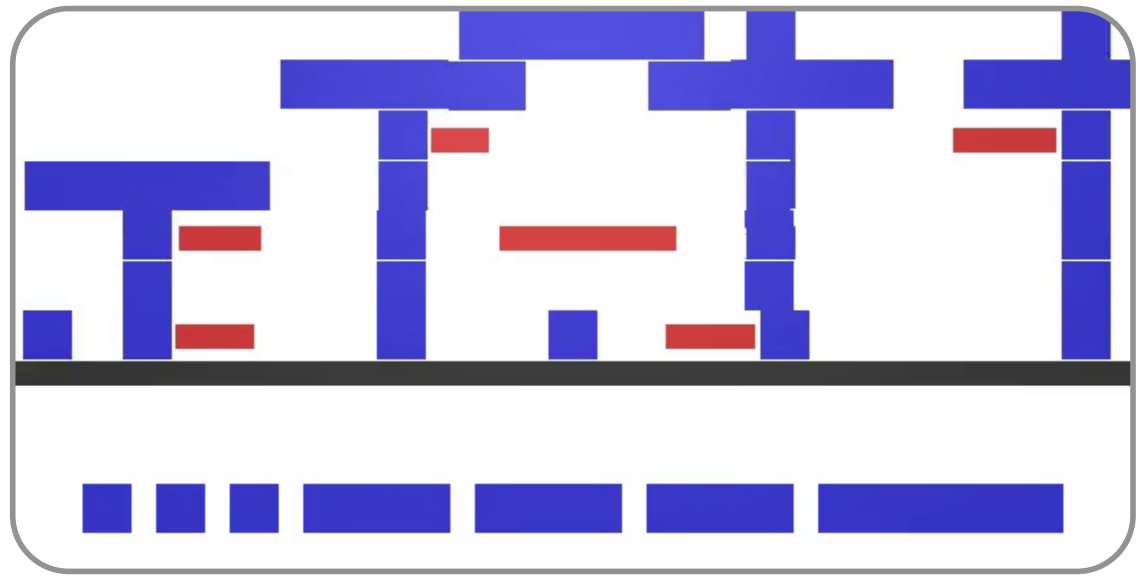
1. **Structured Computation**: the way that individual computations or functions are composed into more complex structures.

   Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (2017, ICLR)



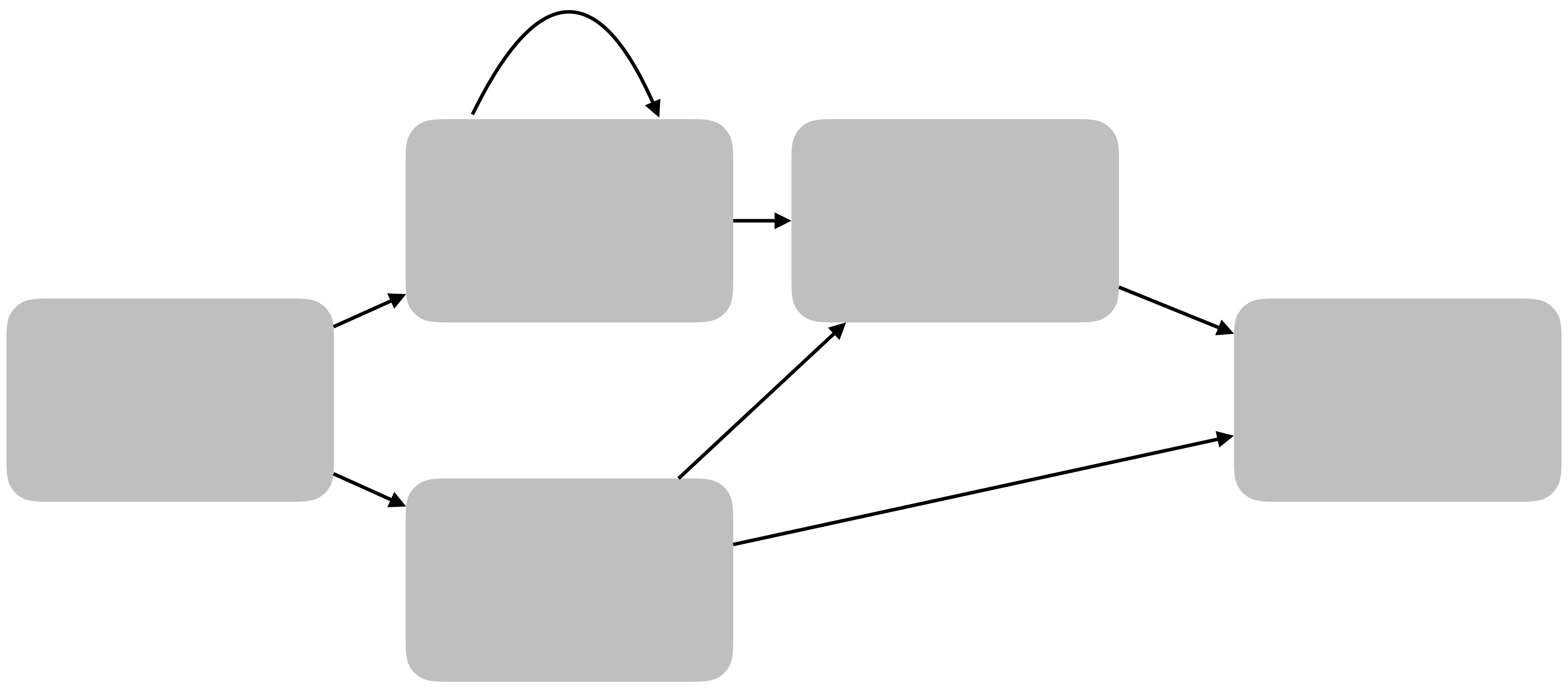
2. **Structured Representation**: the format of the data that computations are performed over.

   Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (2018, arXiv)
   Bapst, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kohli, Battaglia, & Hamrick (2019, ICML)*

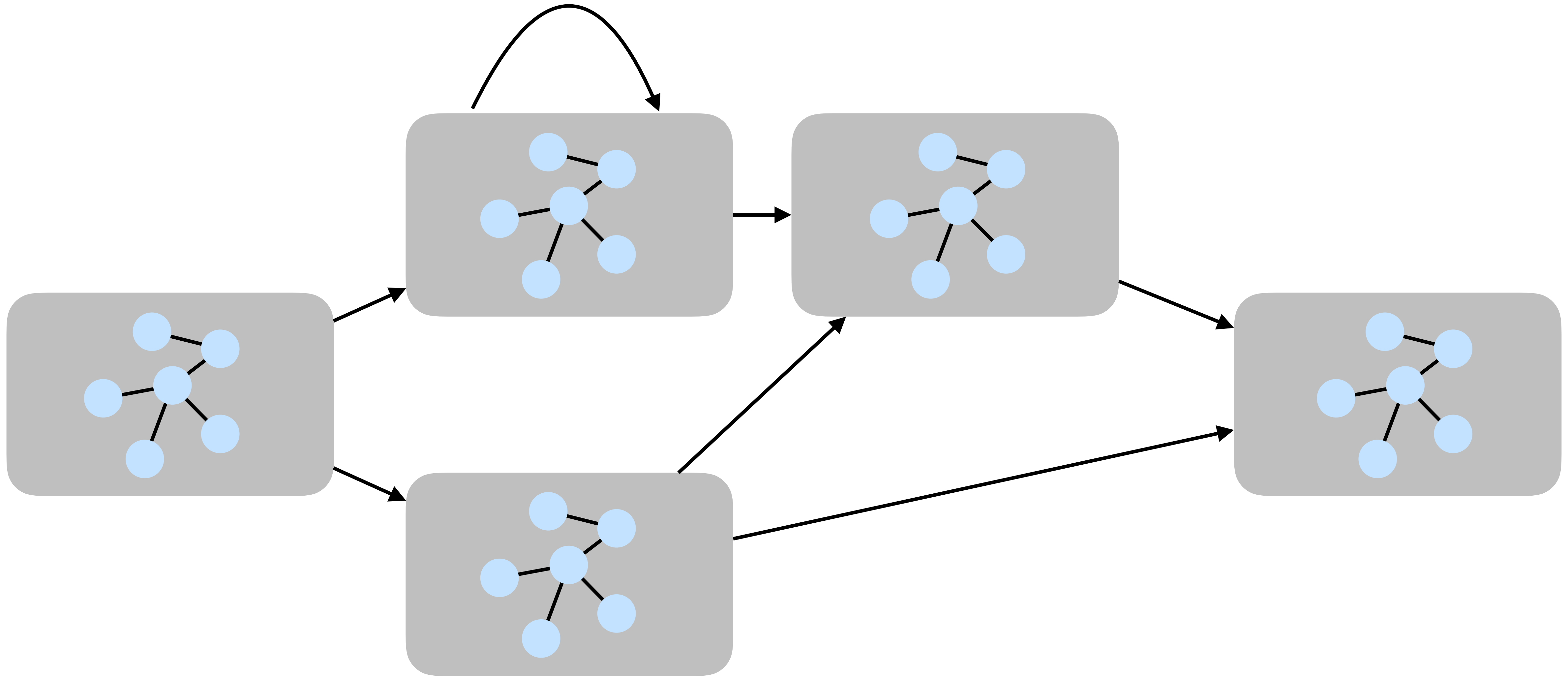


Flexibility, adaptivity and generalization are about ***having the right structure***.

Flexibility, adaptivity and generalization are about ***having the right structure***.

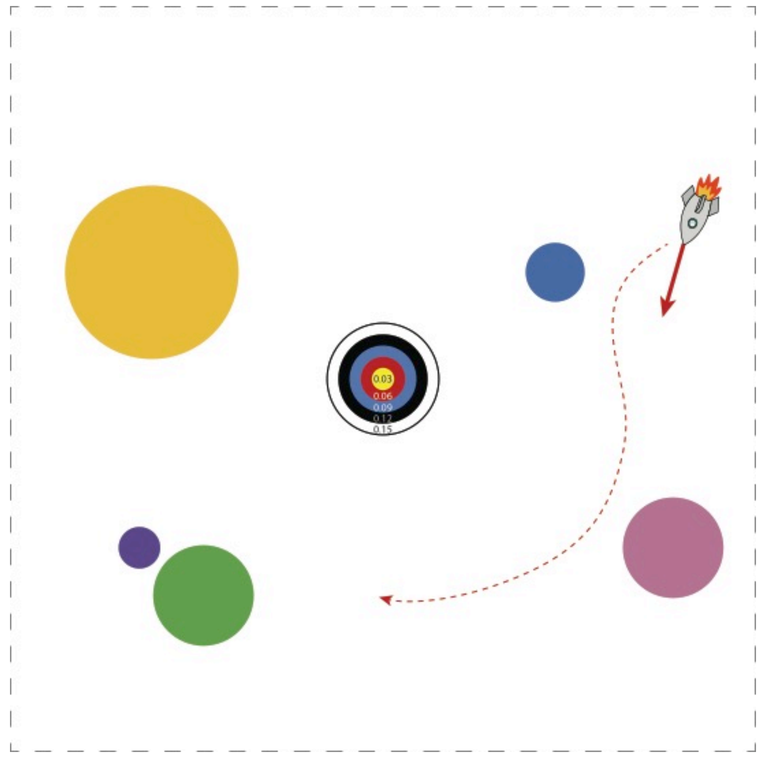


Flexibility, adaptivity and generalization are about ***having the right structure***.

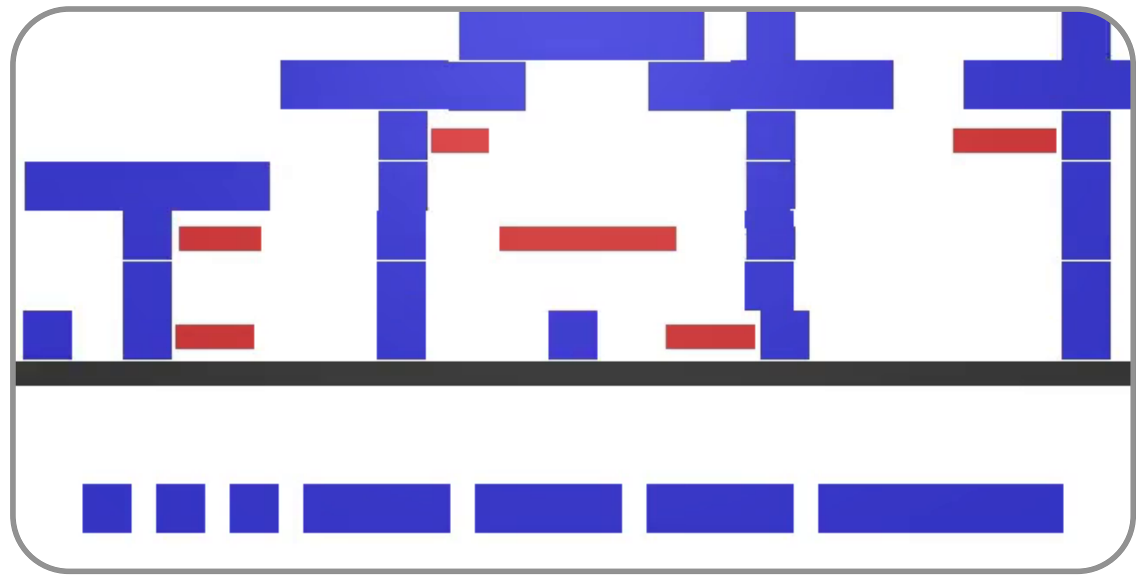


Thanks!

Andy Ballard
Victor Bapst
Peter Battaglia
Hanjun Dai
Carl Doersch
Nicolas Heess
Thomas Kipf
Pushmeet Kohli
Yujia Li
Razvan Pascanu
Tobias Pfaff
David Reichert
Alvaro Sanchez-Gonzalez
Kim Stachenfeld
Oriol Vinyals
Theo Weber

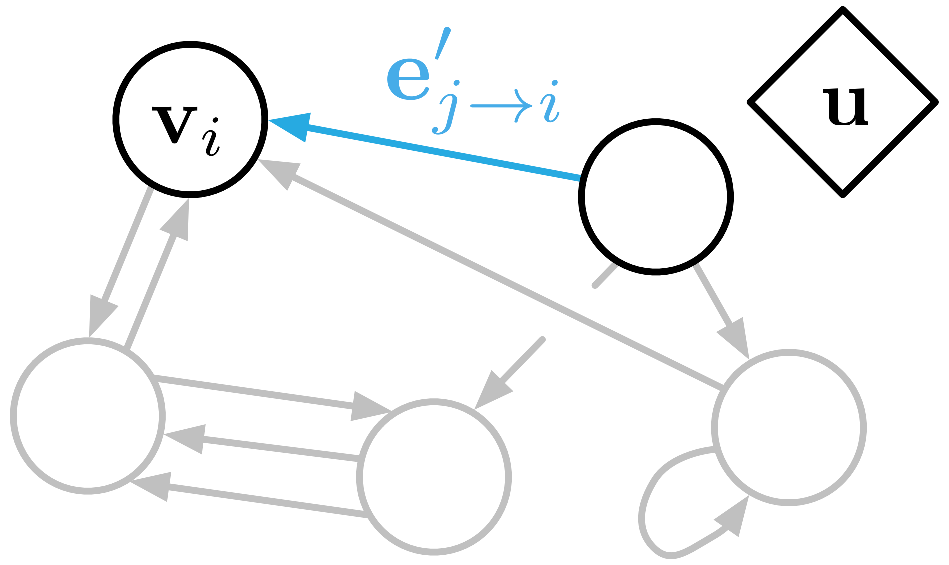


Hamrick, Ballard, Pascanu, Vinyals, Heess, & Battaglia (*ICLR* 2017)



Bapst*, Sanchez-Gonzalez*, Doersch, Stachenfeld, Kolhli, Battaglia, & Hamrick (*ICML* 2019)

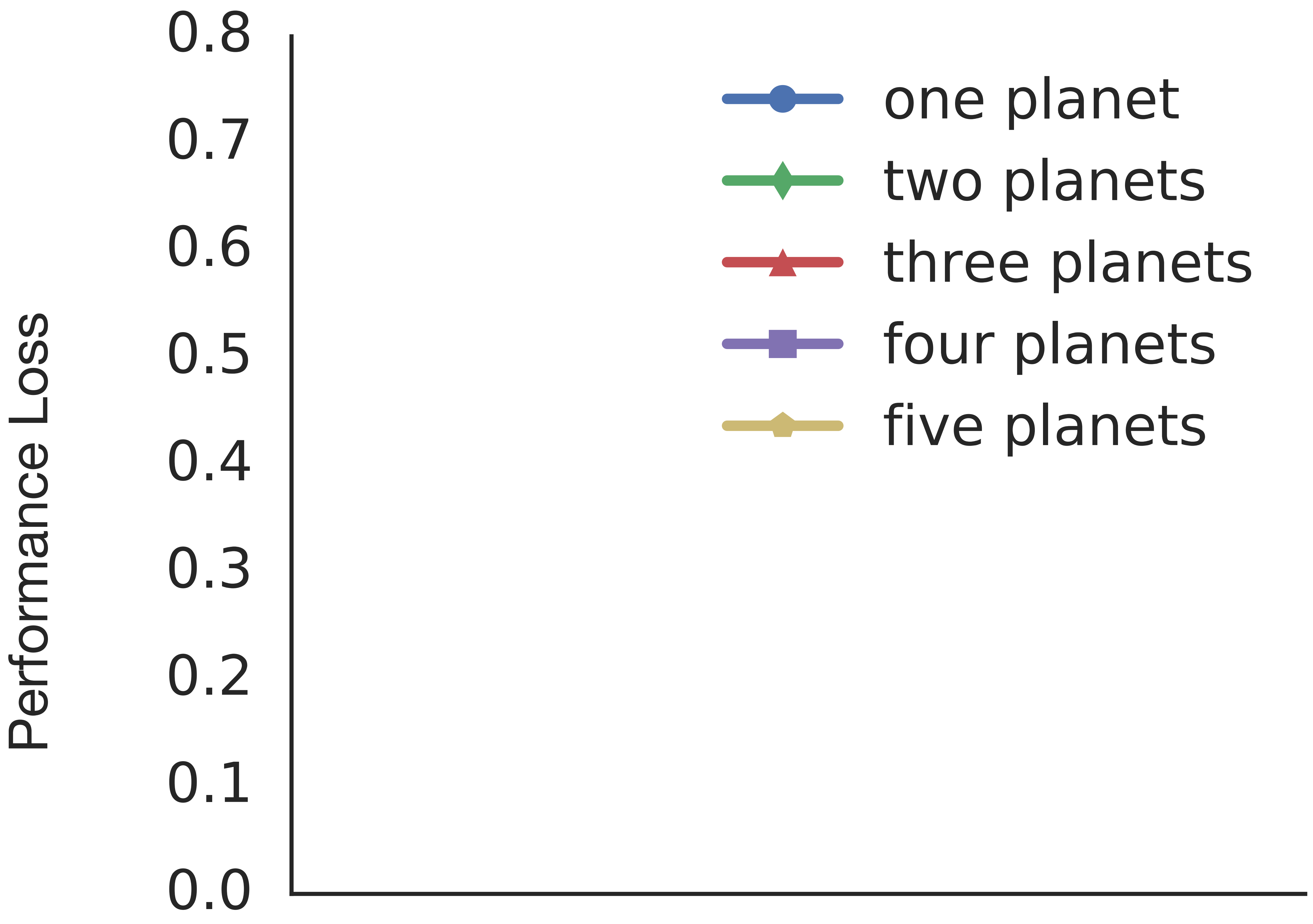
**equal contribution*



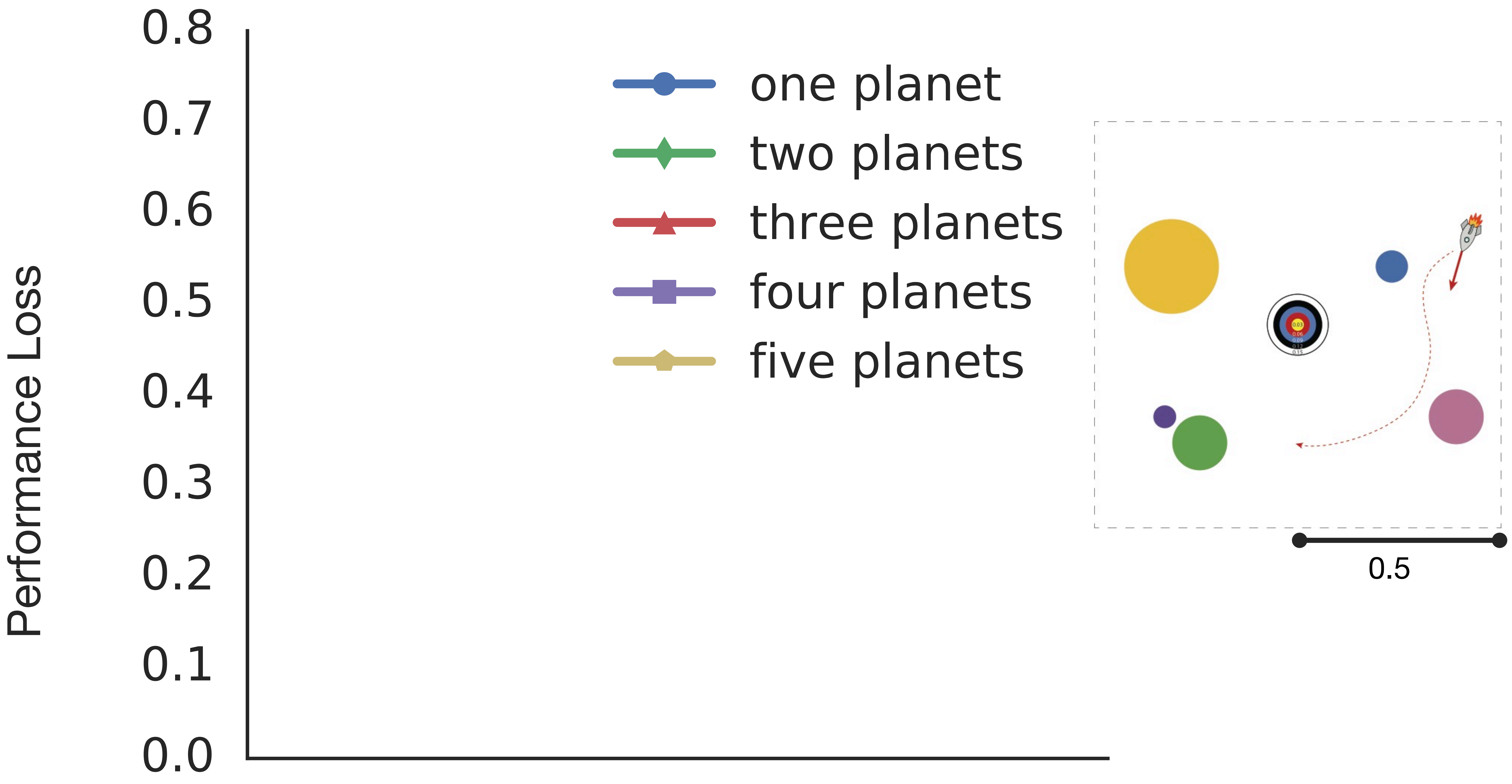
Battaglia, Hamrick, Bapst, Sanchez-Gonzalez, Zambaldi, et al. (*arXiv* 2018)

Extra Slides

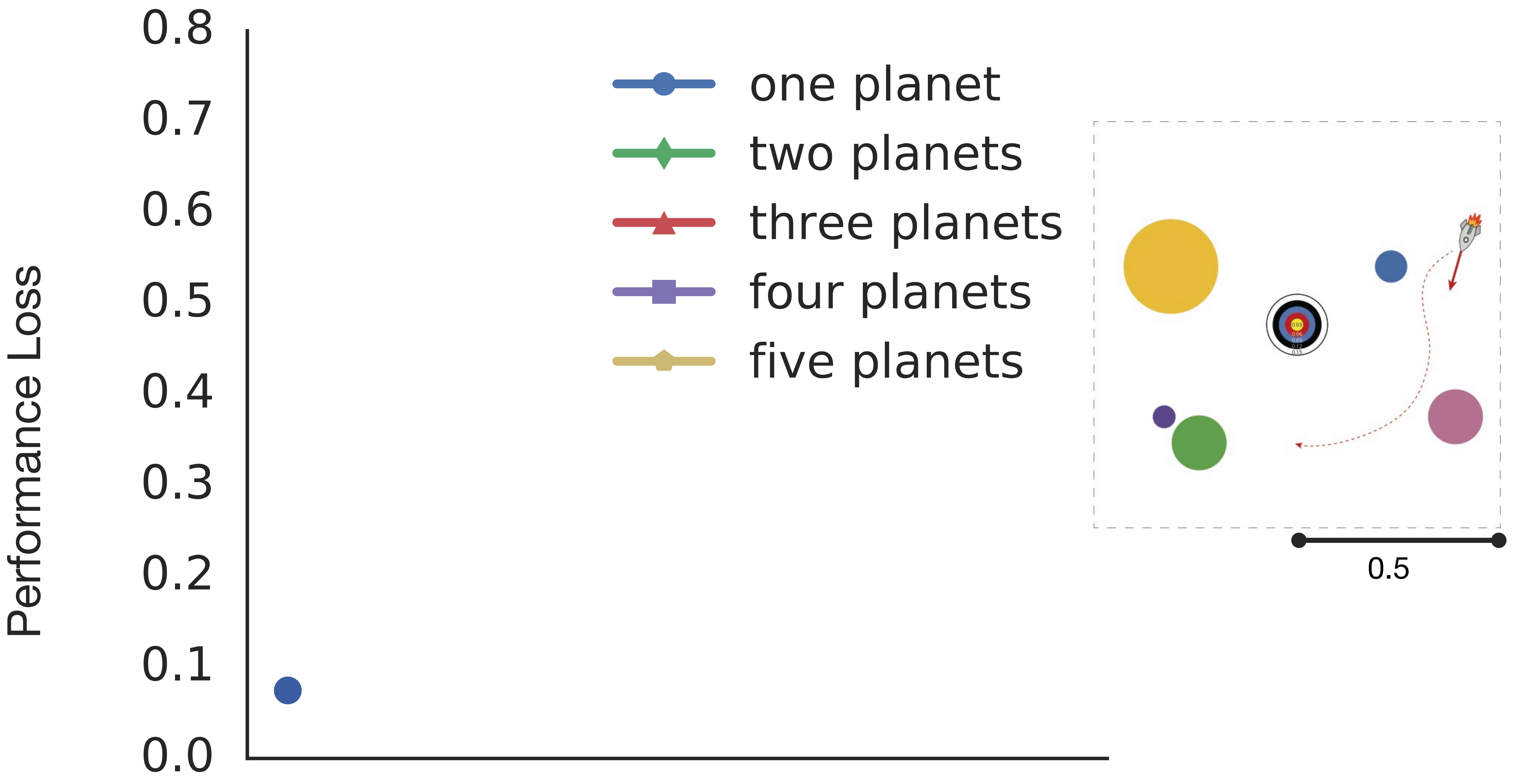
Reactive Controller Results



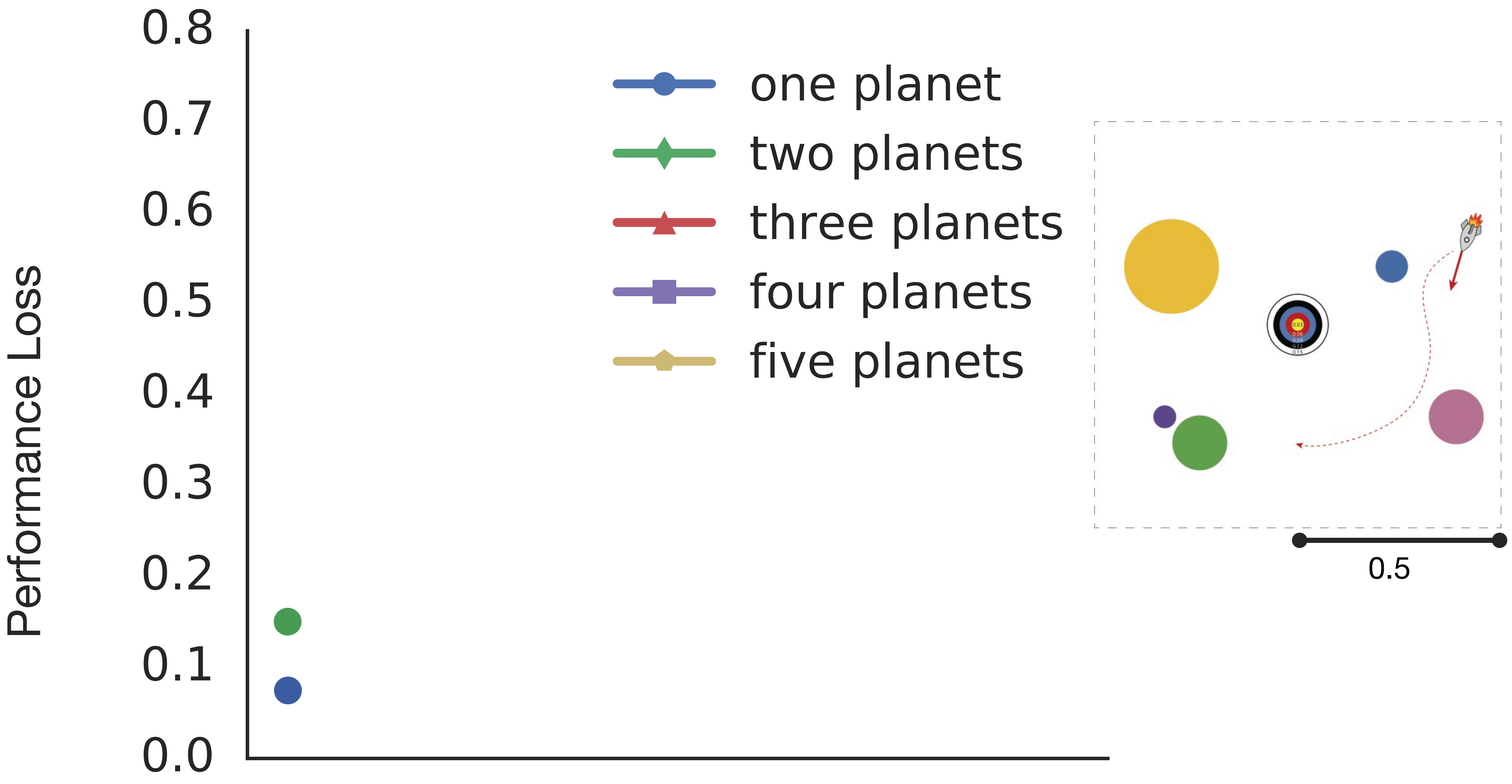
Reactive Controller Results



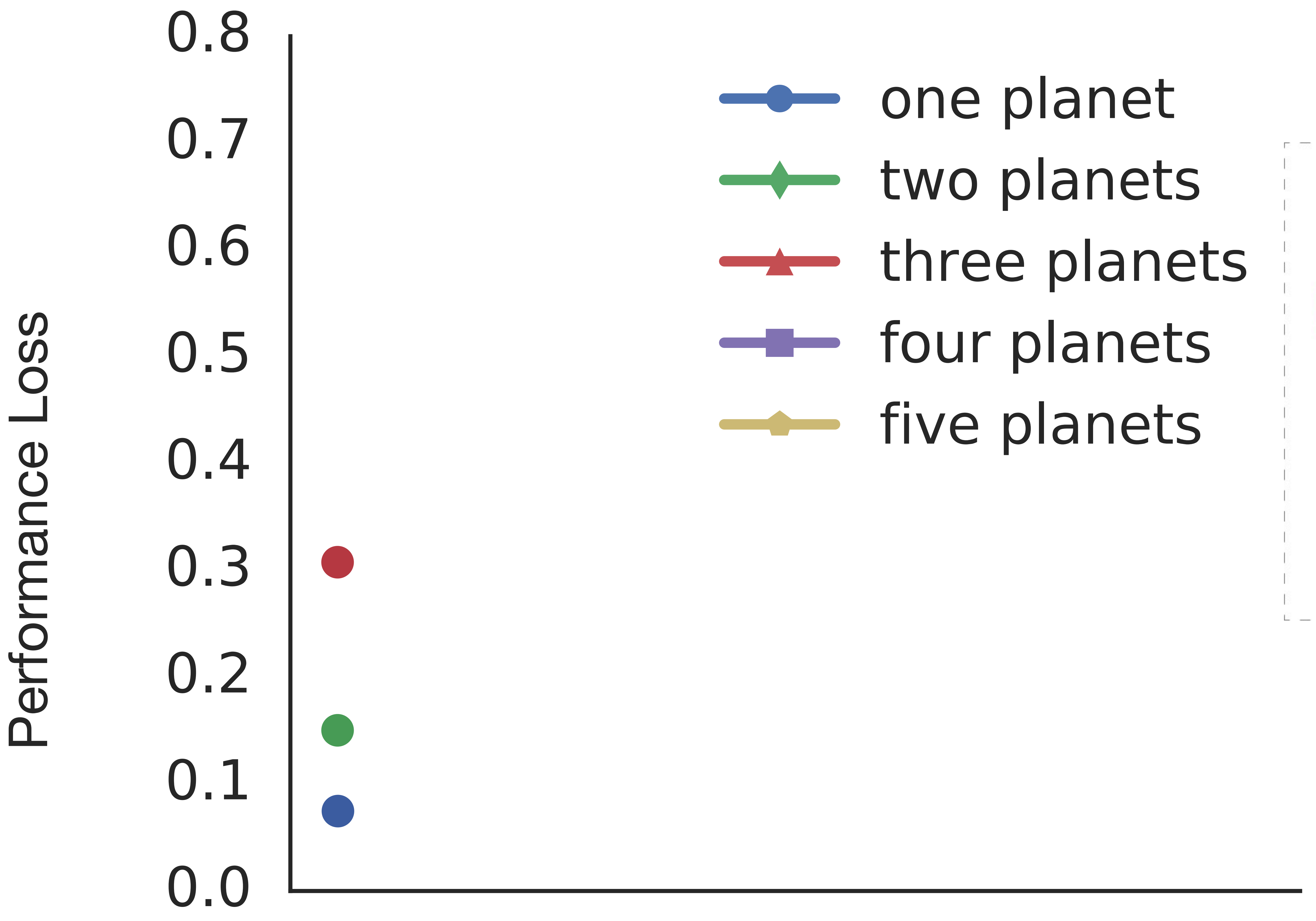
Reactive Controller Results

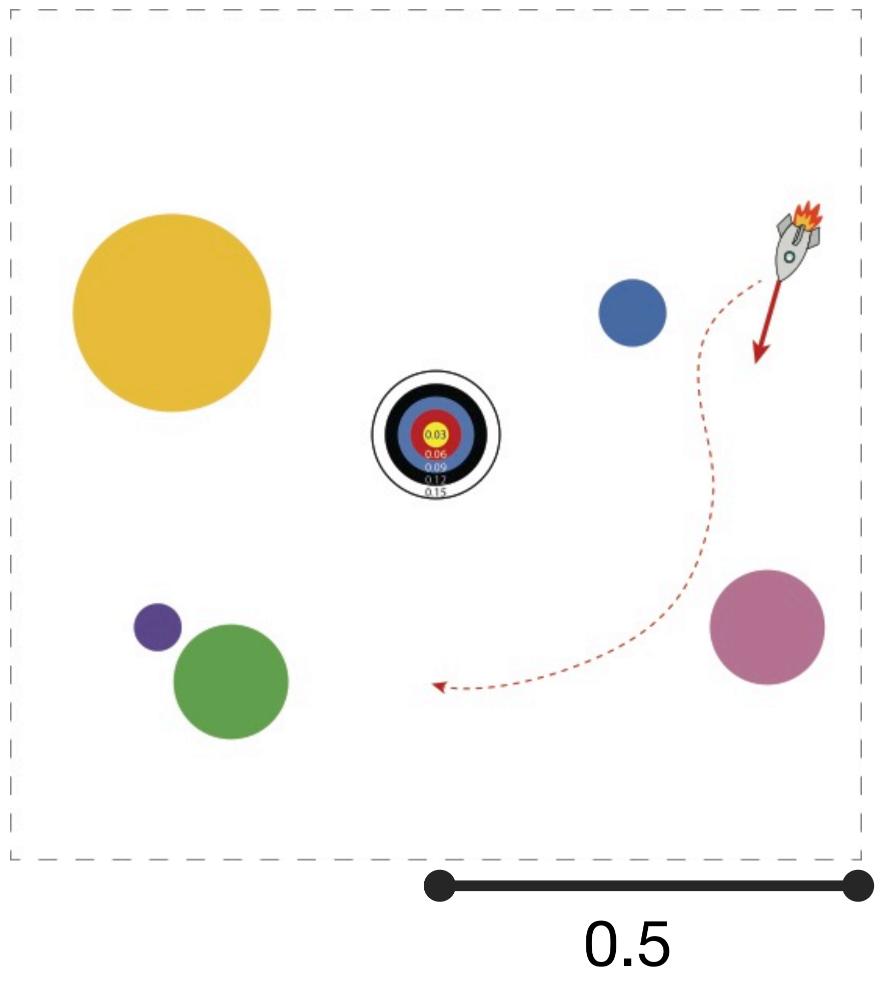


Reactive Controller Results

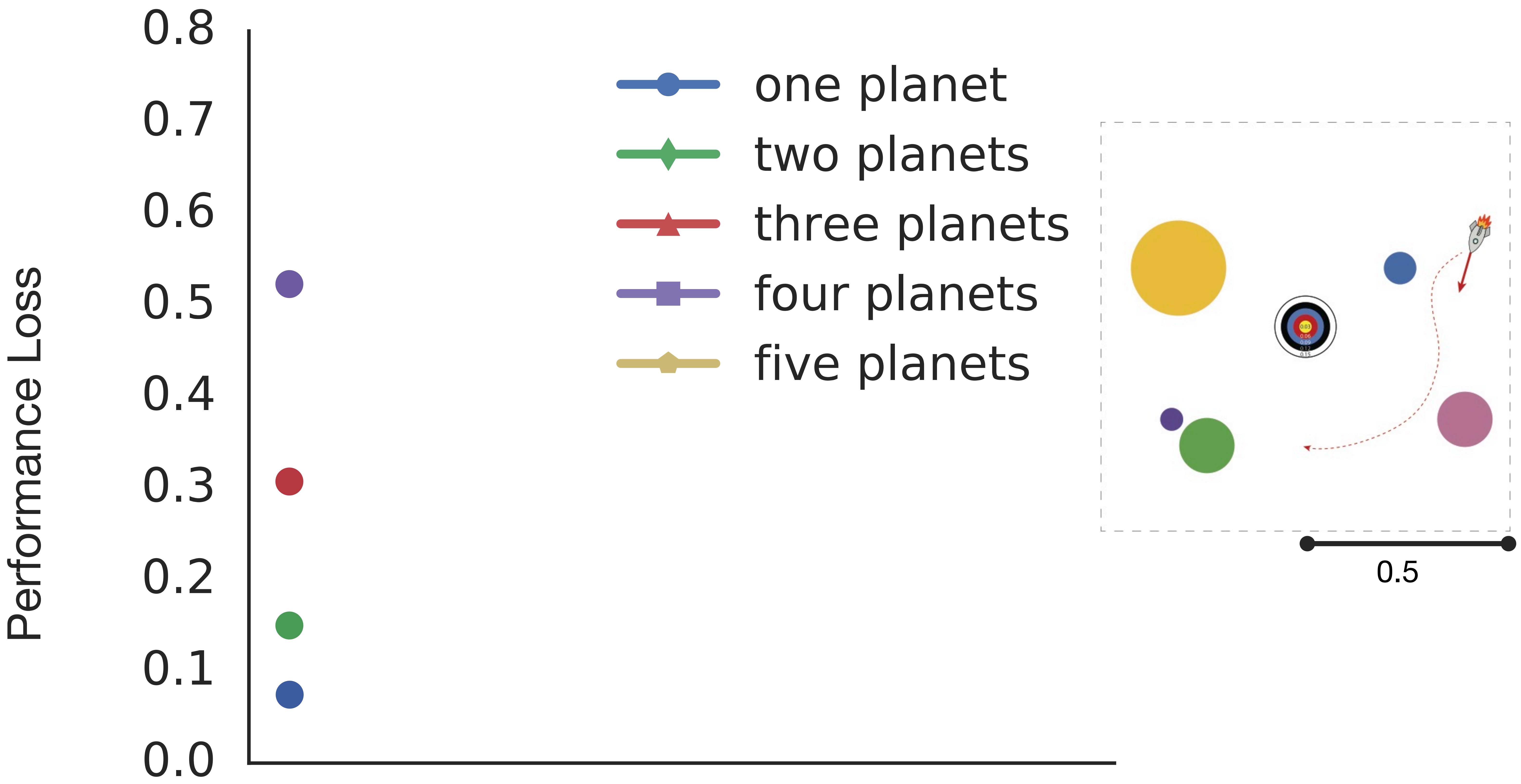


Reactive Controller Results

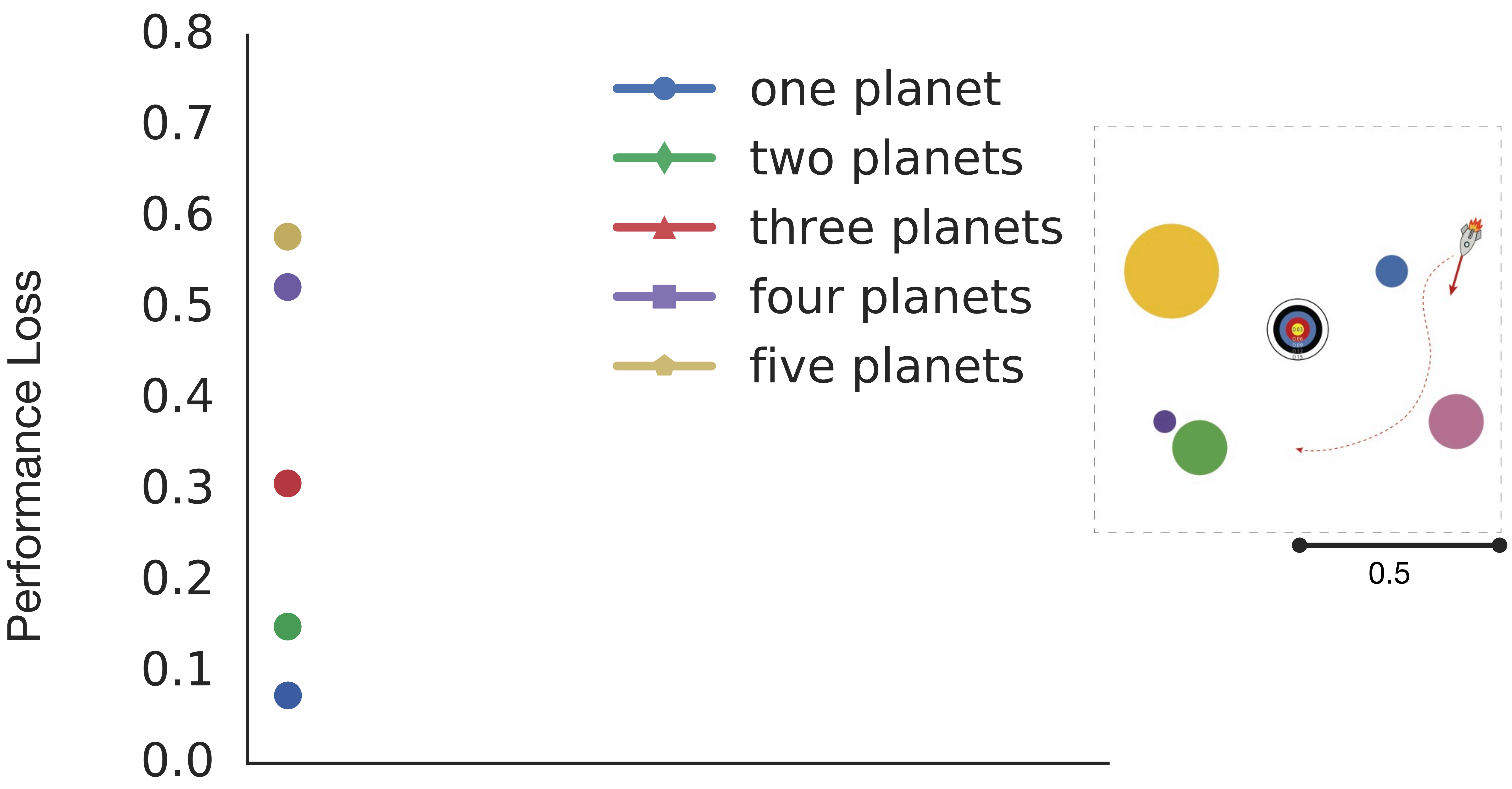




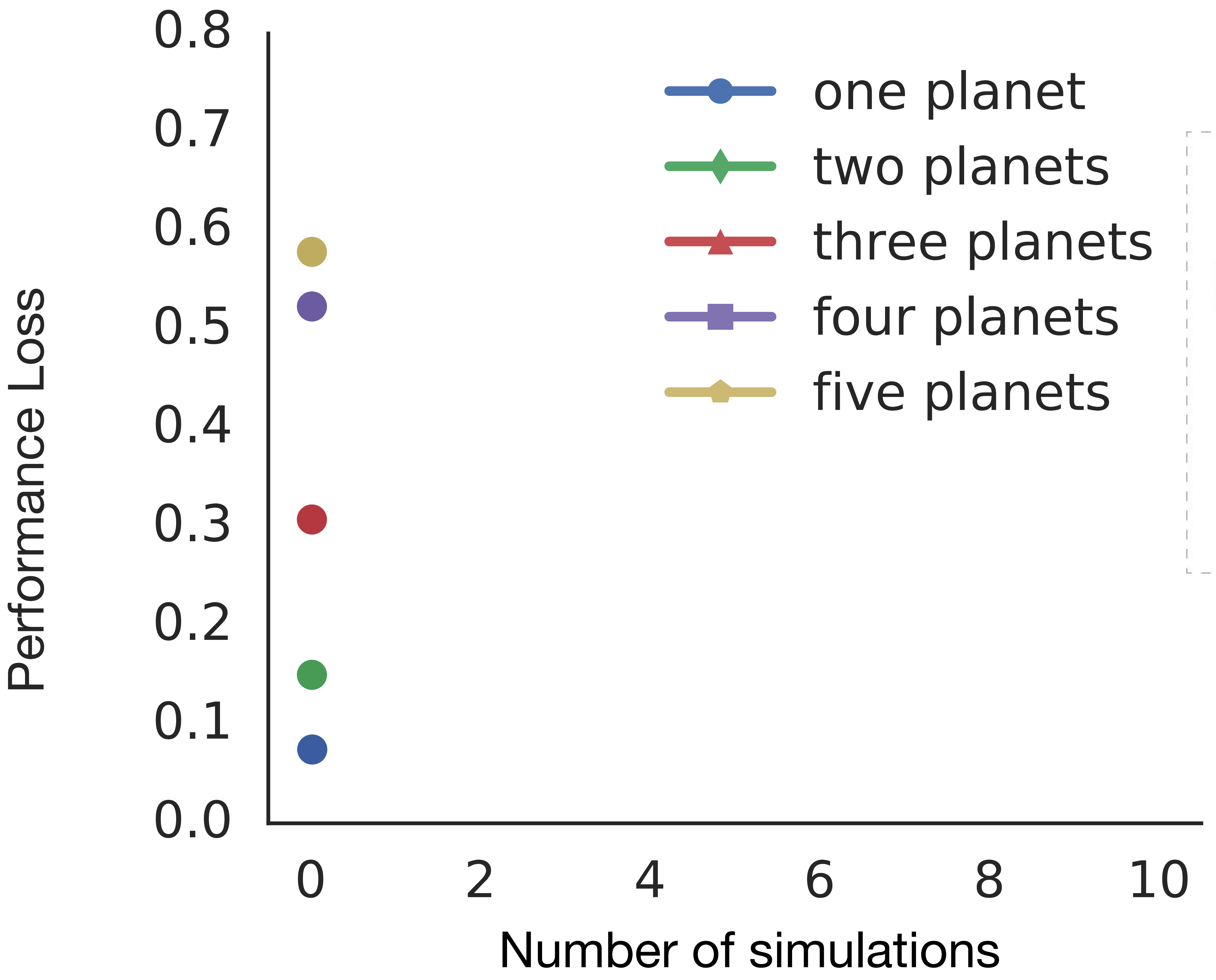
Reactive Controller Results

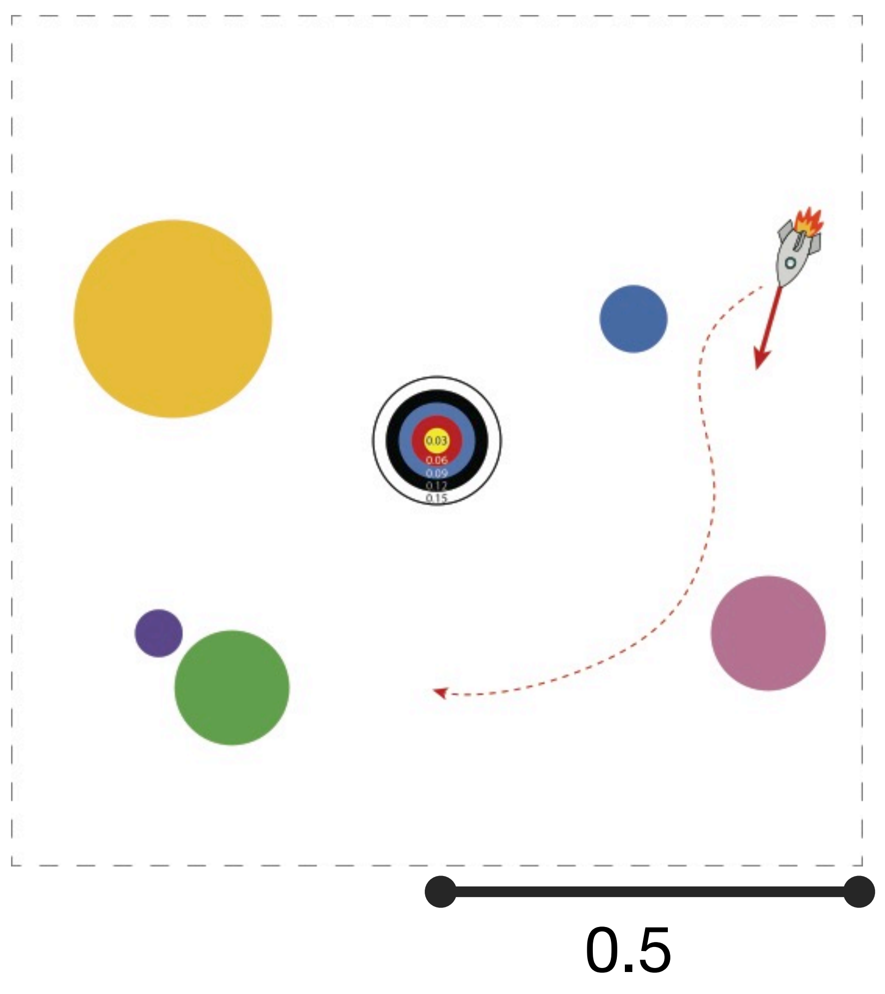


Reactive Controller Results

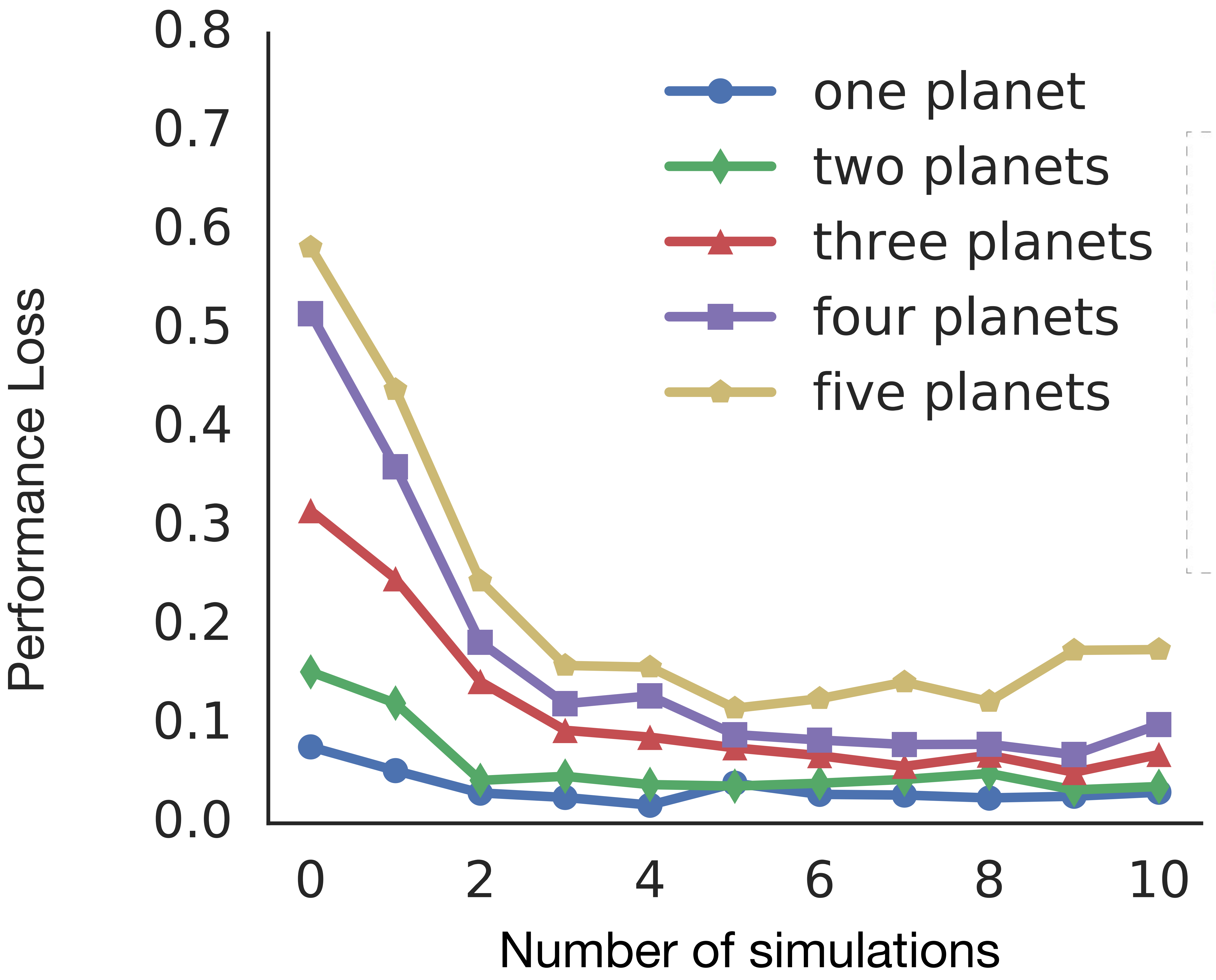


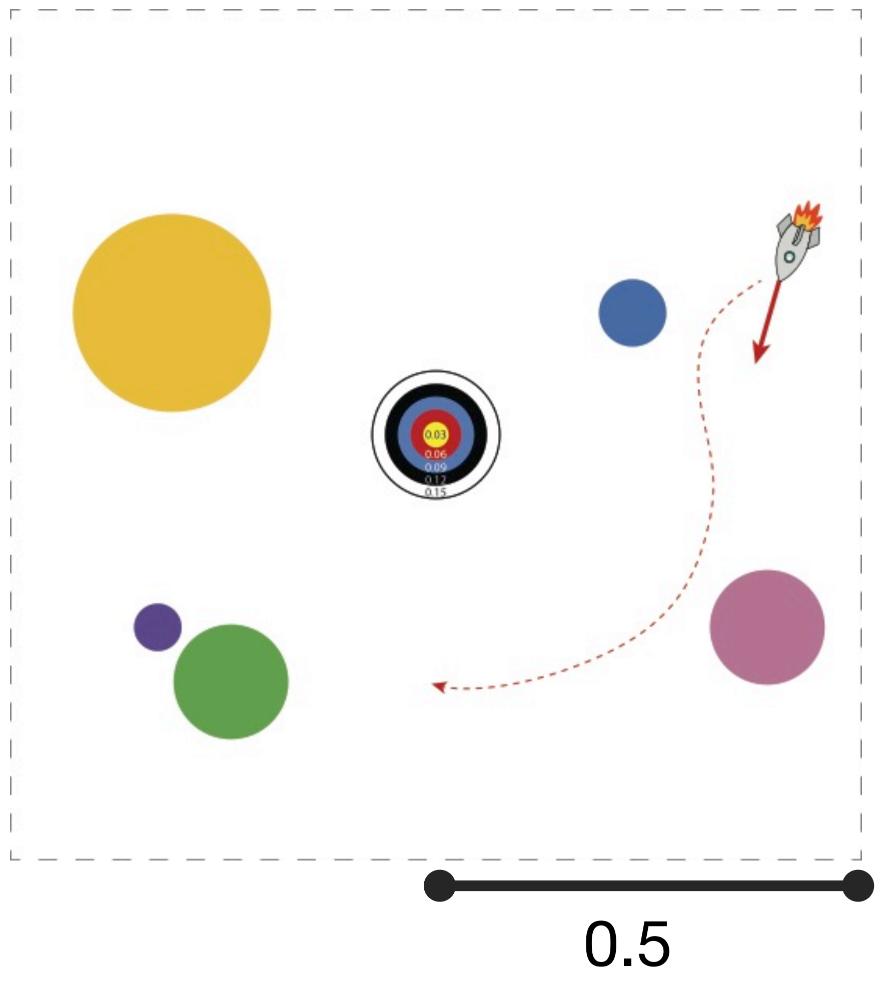
Iterative Controller Results





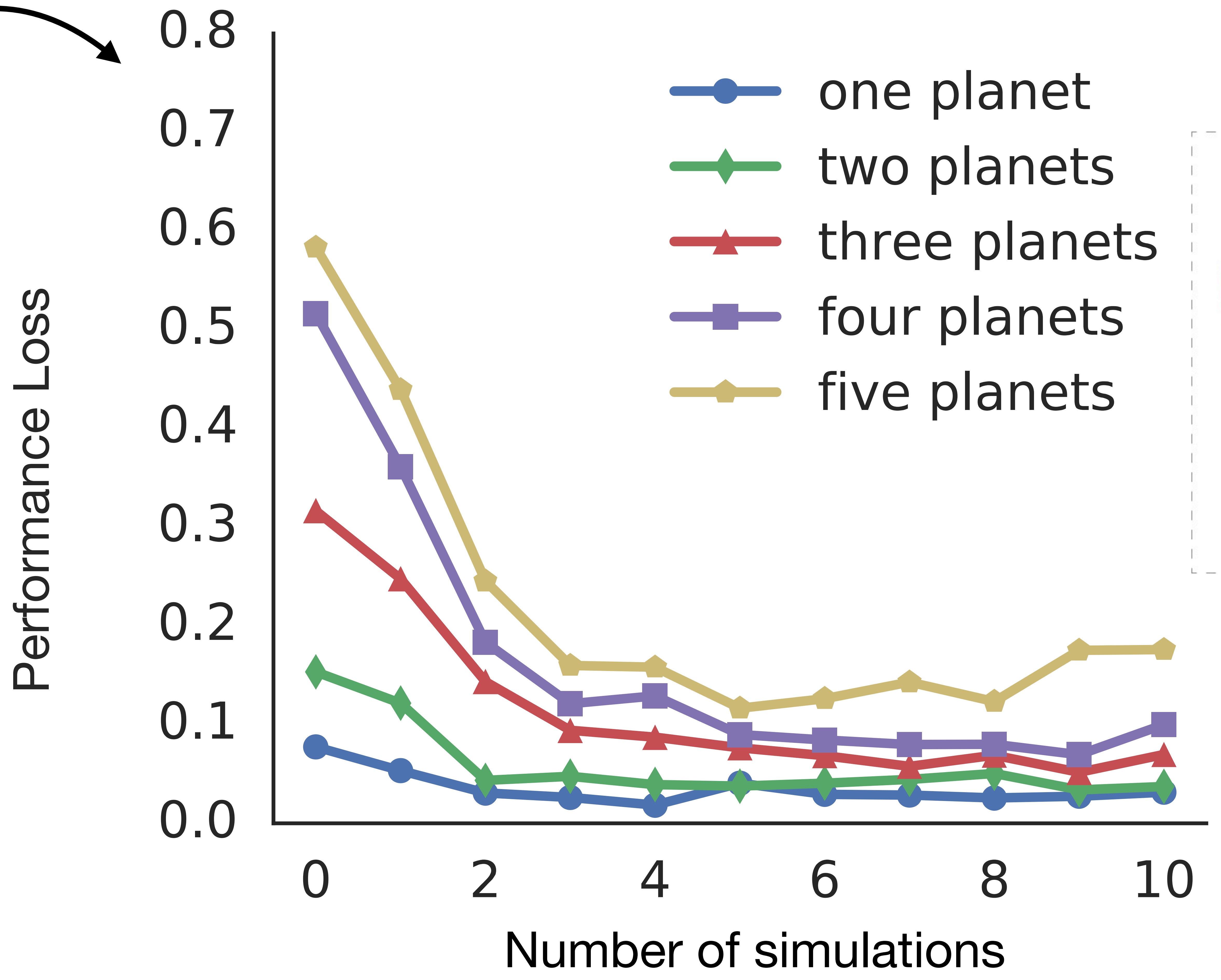
Iterative Controller Results

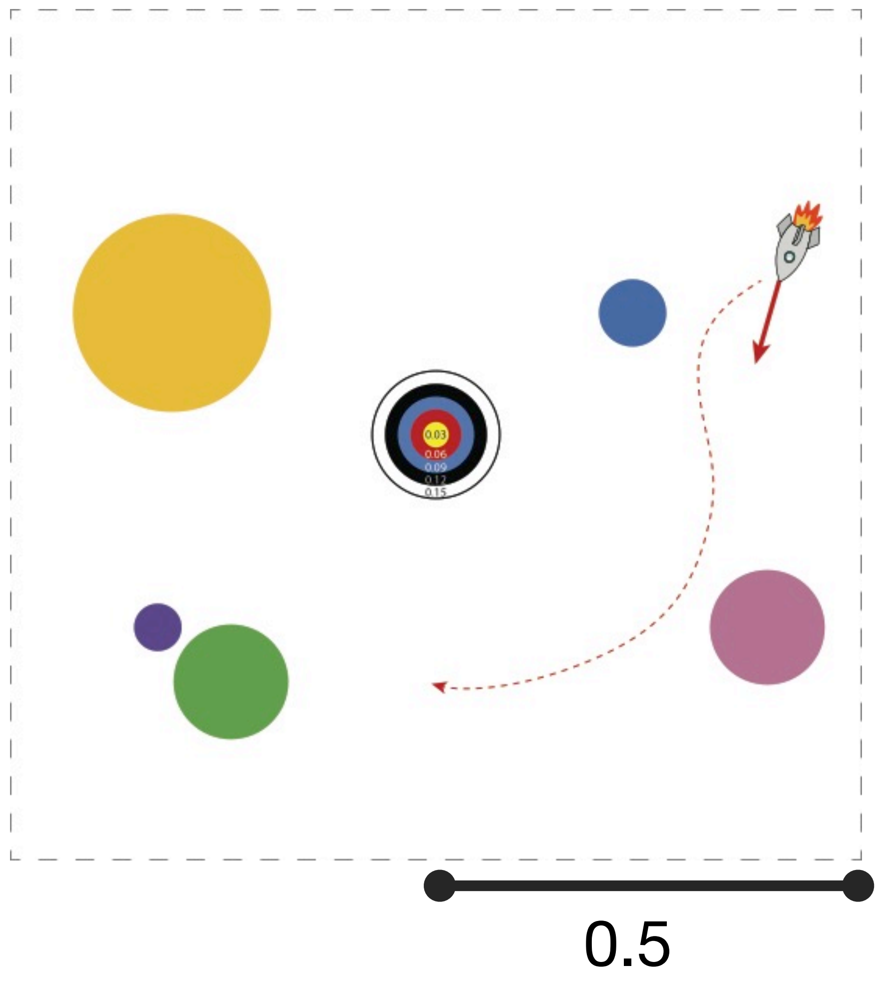




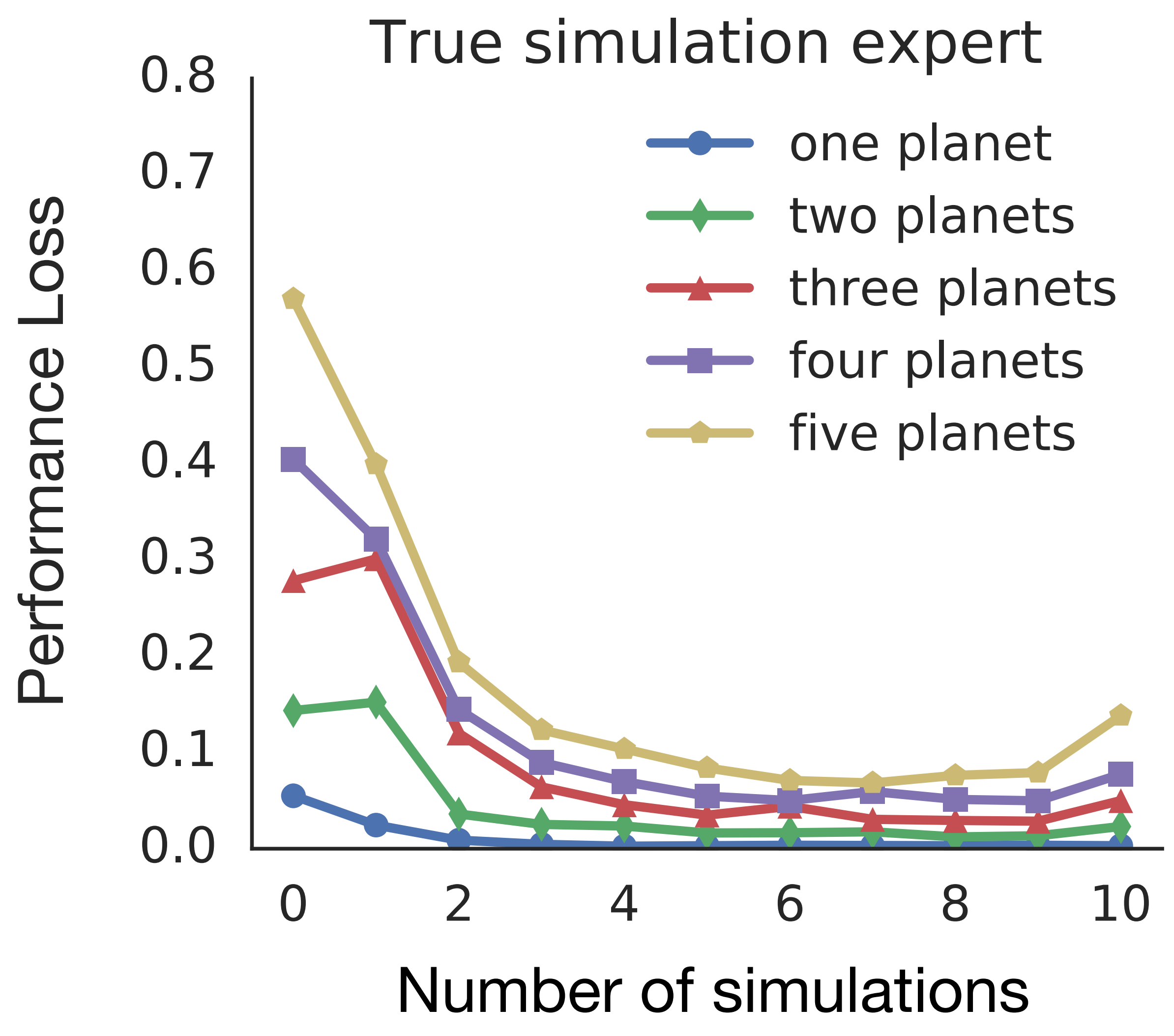
Iterative Controller Results

This is using a learned expert, which is trained simultaneously with the agent

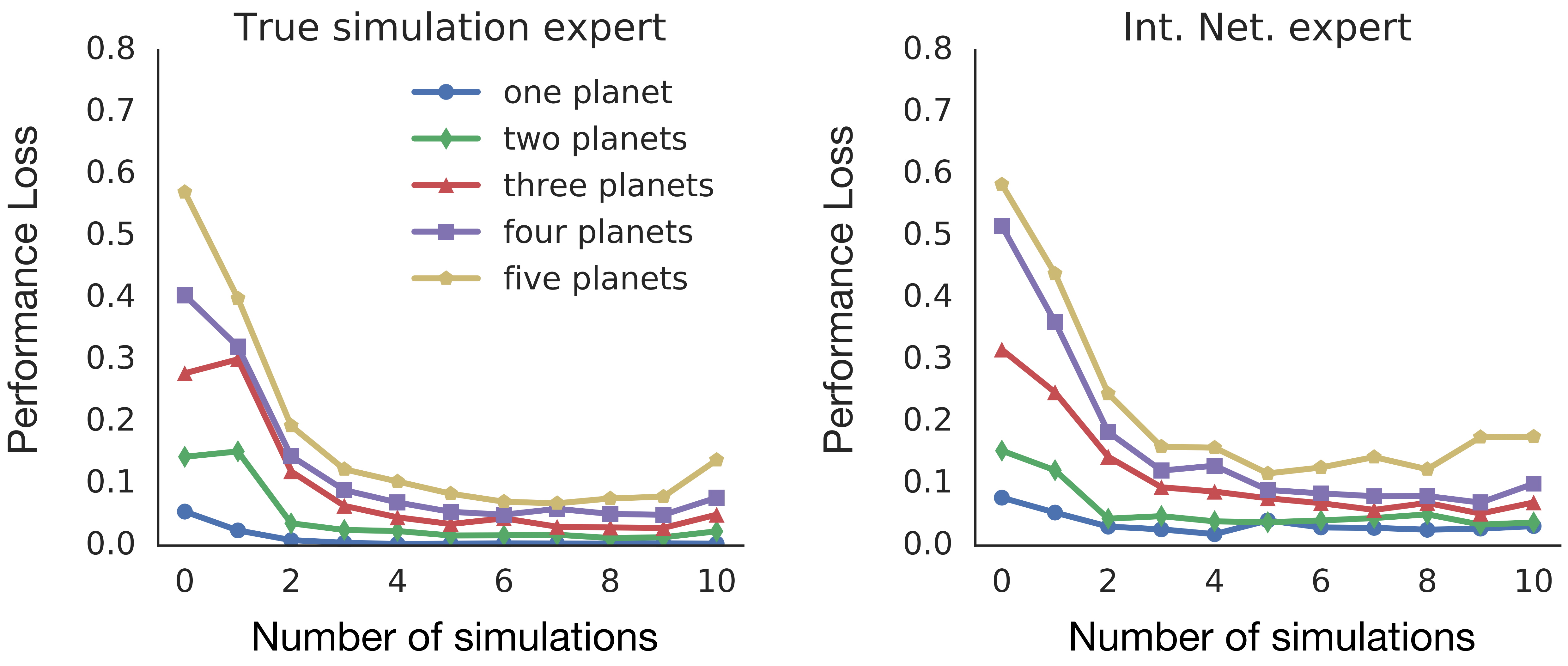




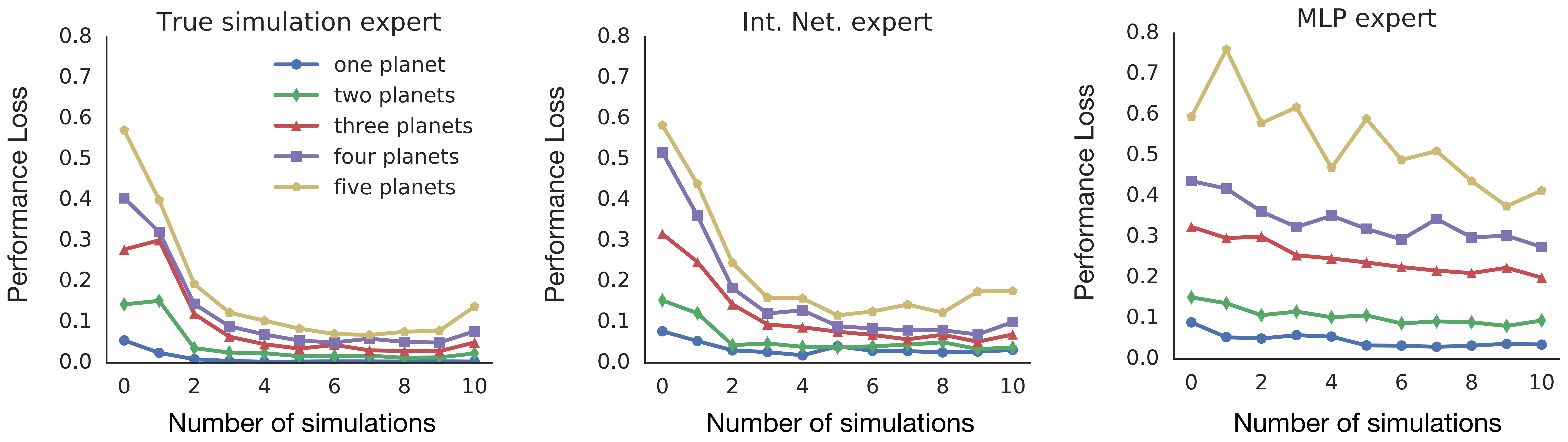
Iterative Controller Results



Iterative Controller Results



Iterative Controller Results

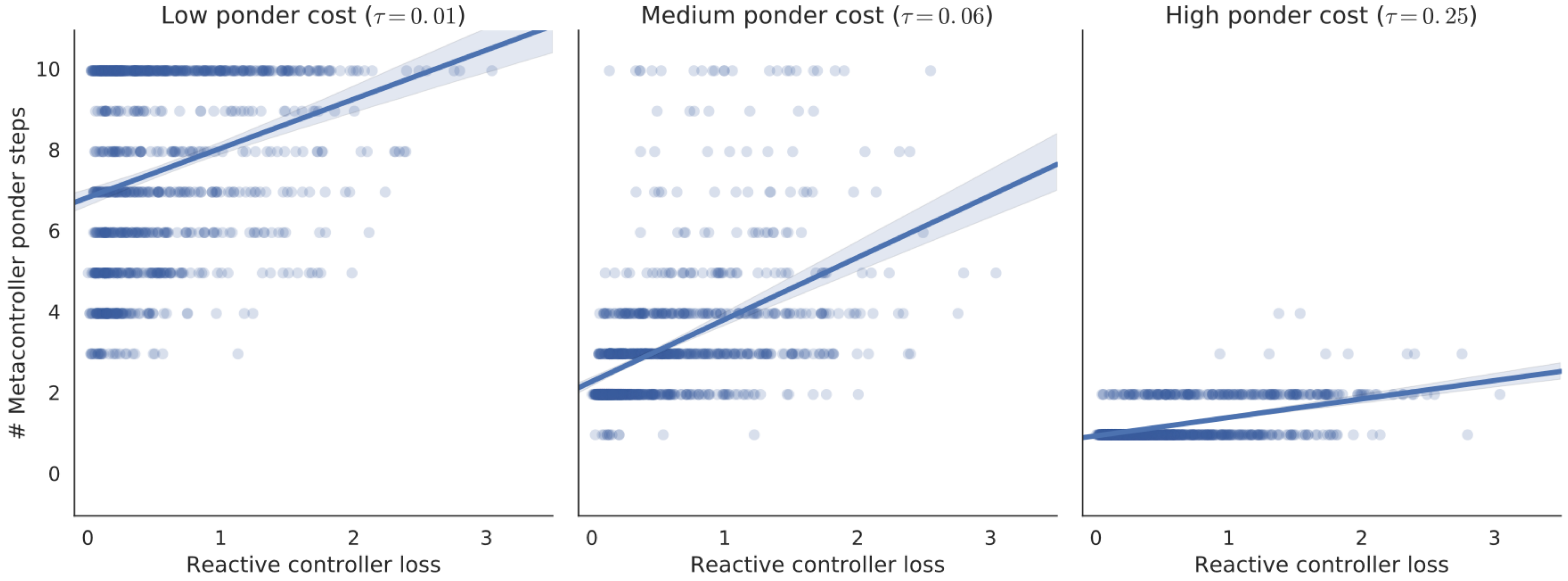


Metacontroller Results

Effect of difficulty (IN expert)

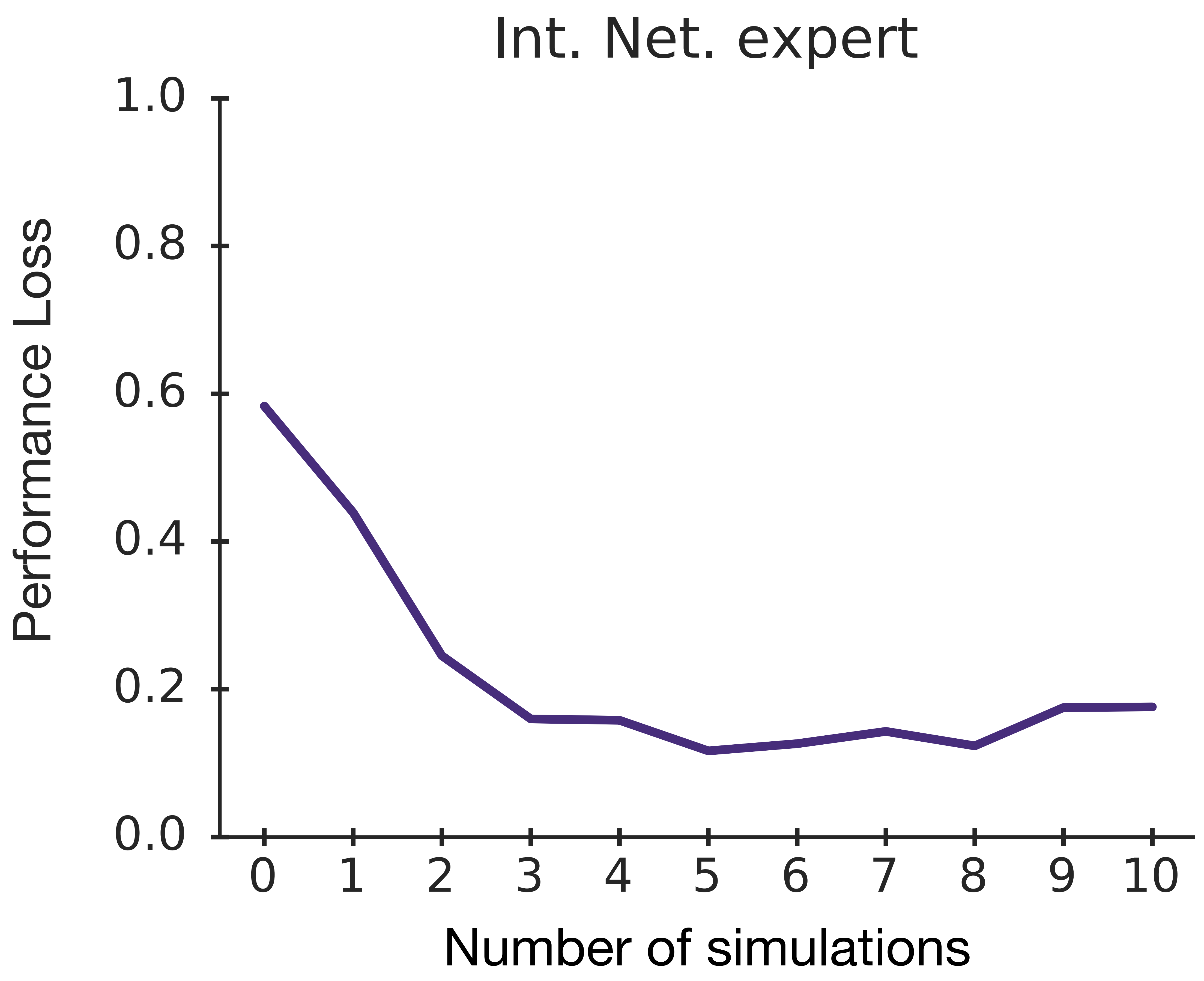
Metacontroller Results

Effect of difficulty (IN expert)



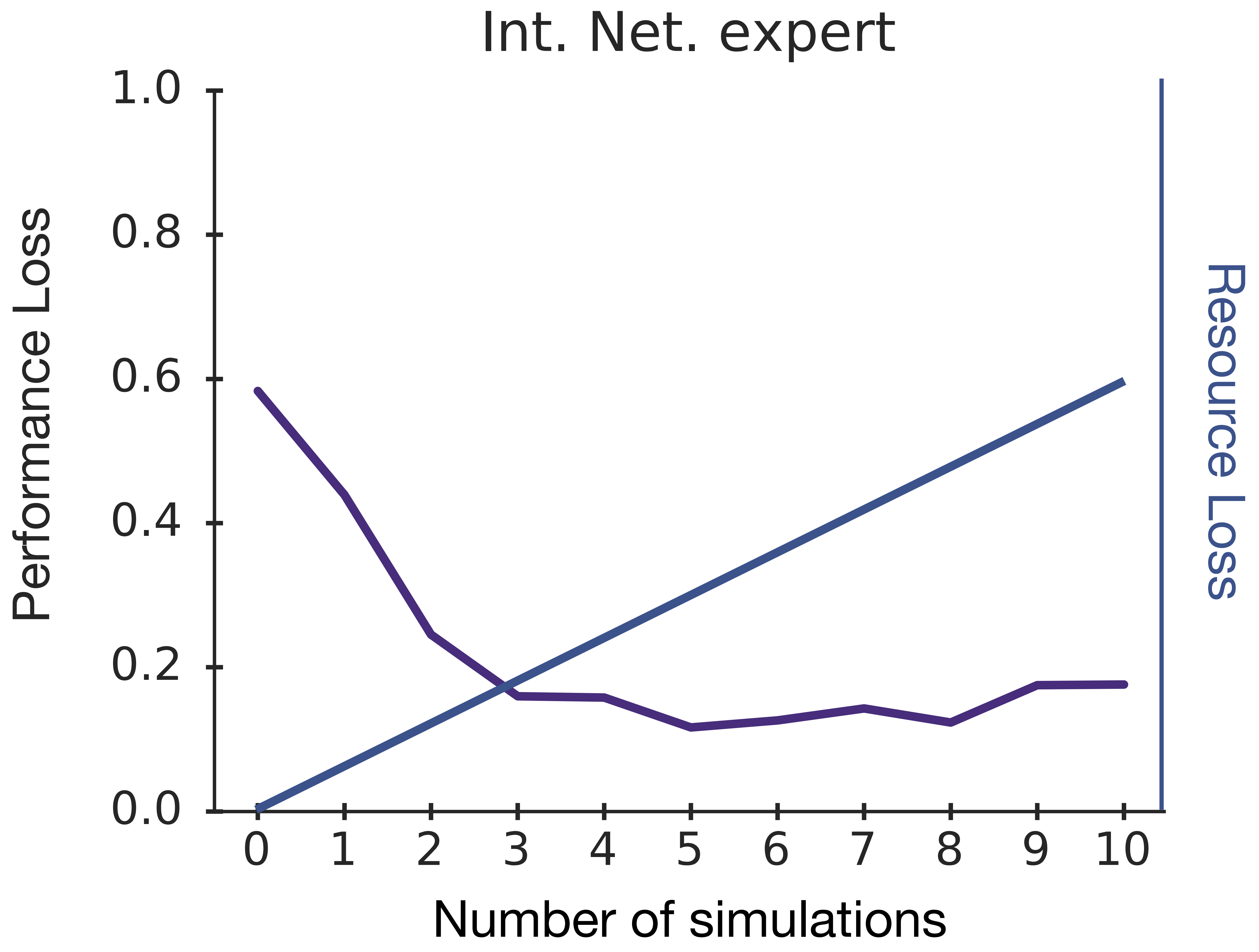
Metacontroller Results

Total cost (five planets)



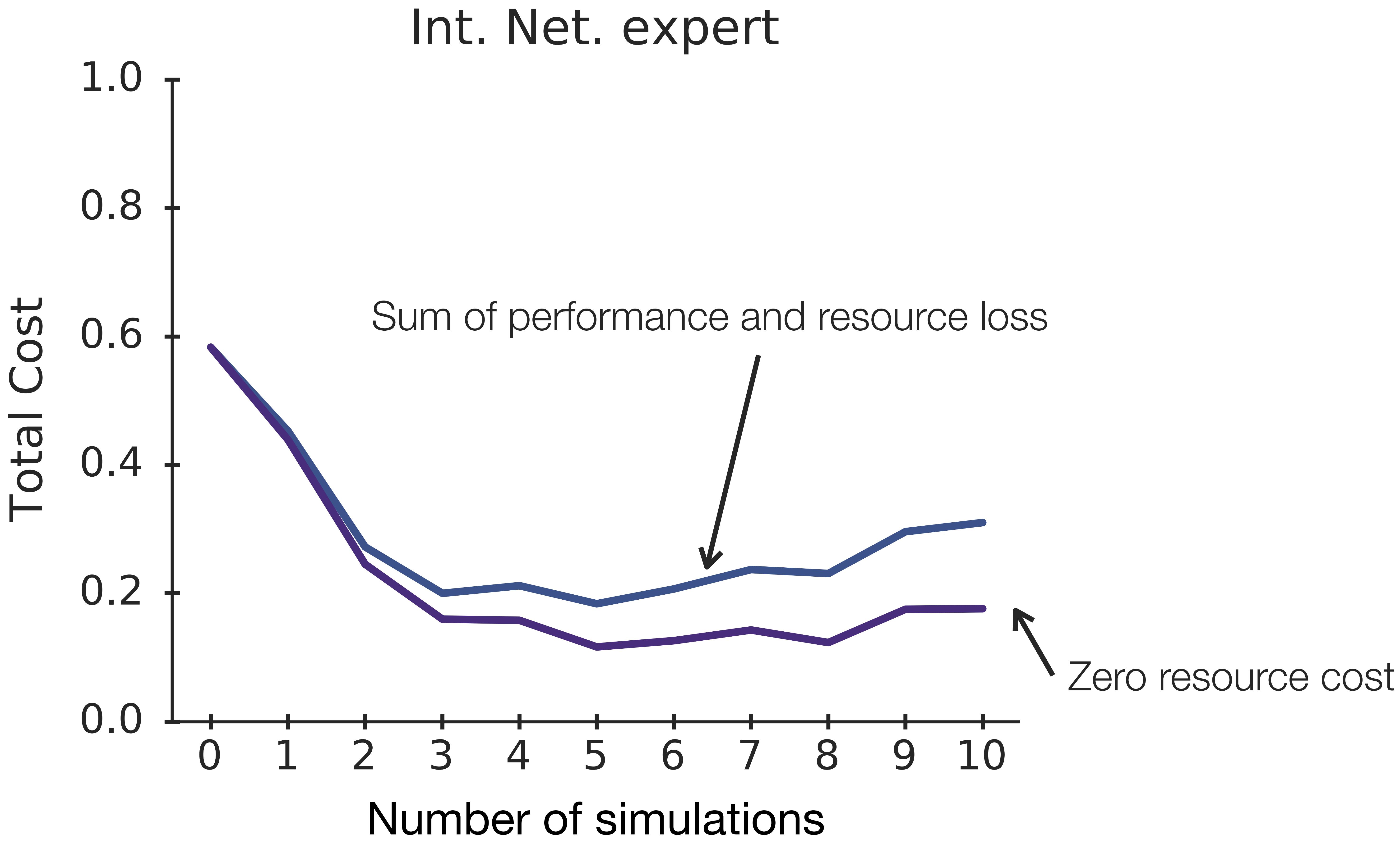
Metacontroller Results

Total cost (five planets)



Metacontroller Results

Total cost (five planets)

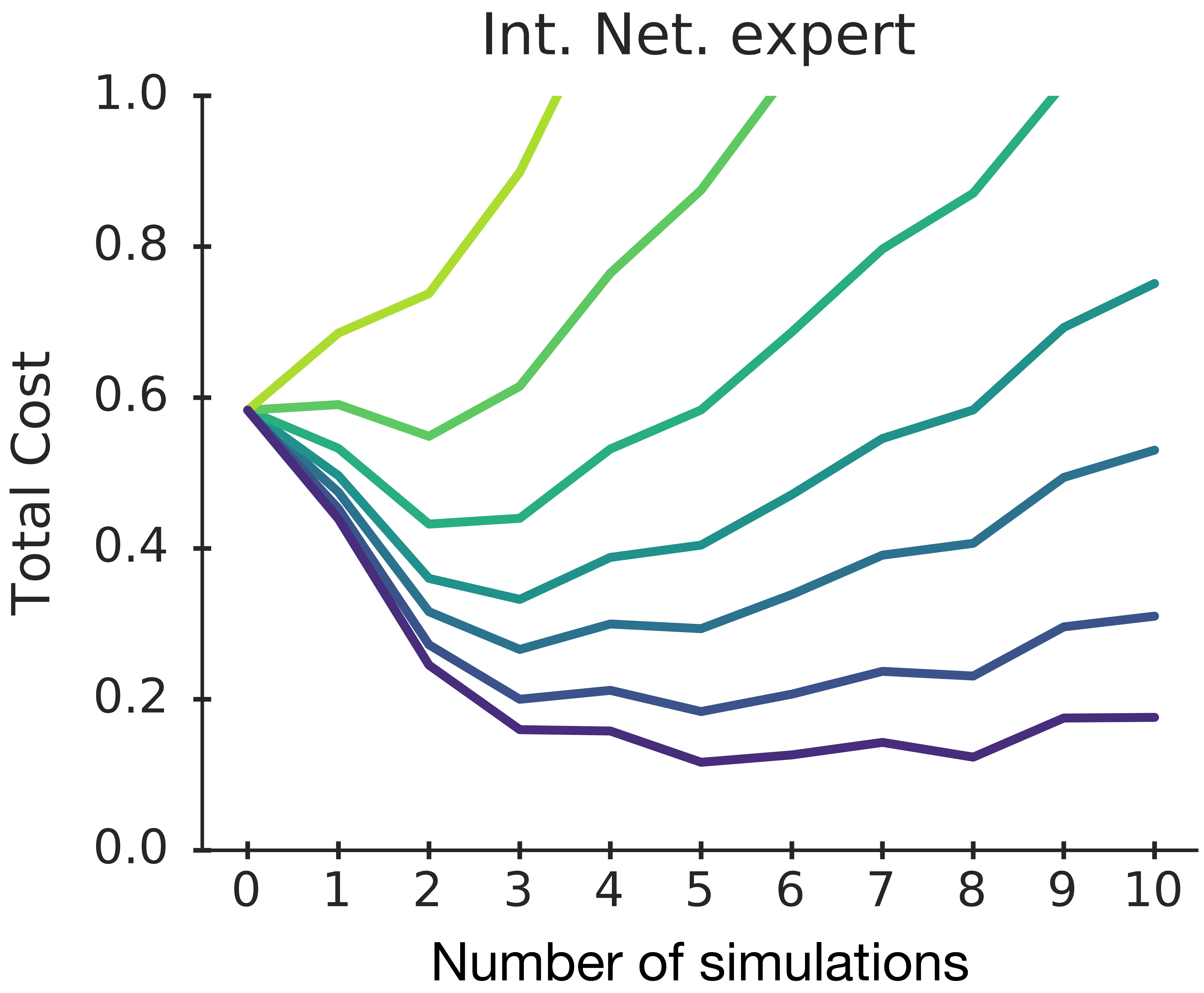


Sum of performance and resource loss

Zero resource cost

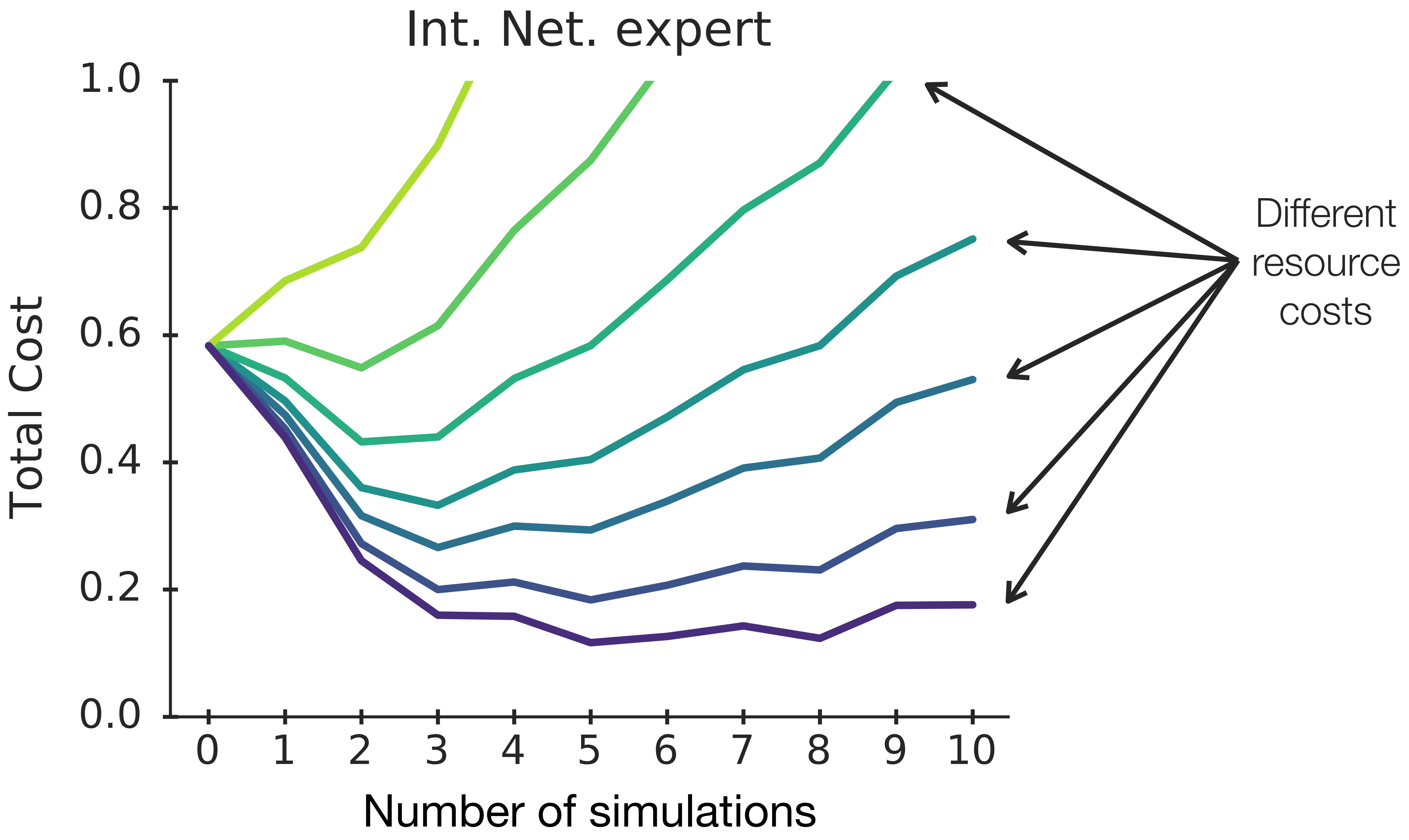
Metacontroller Results

Total cost (five planets)



Metacontroller Results

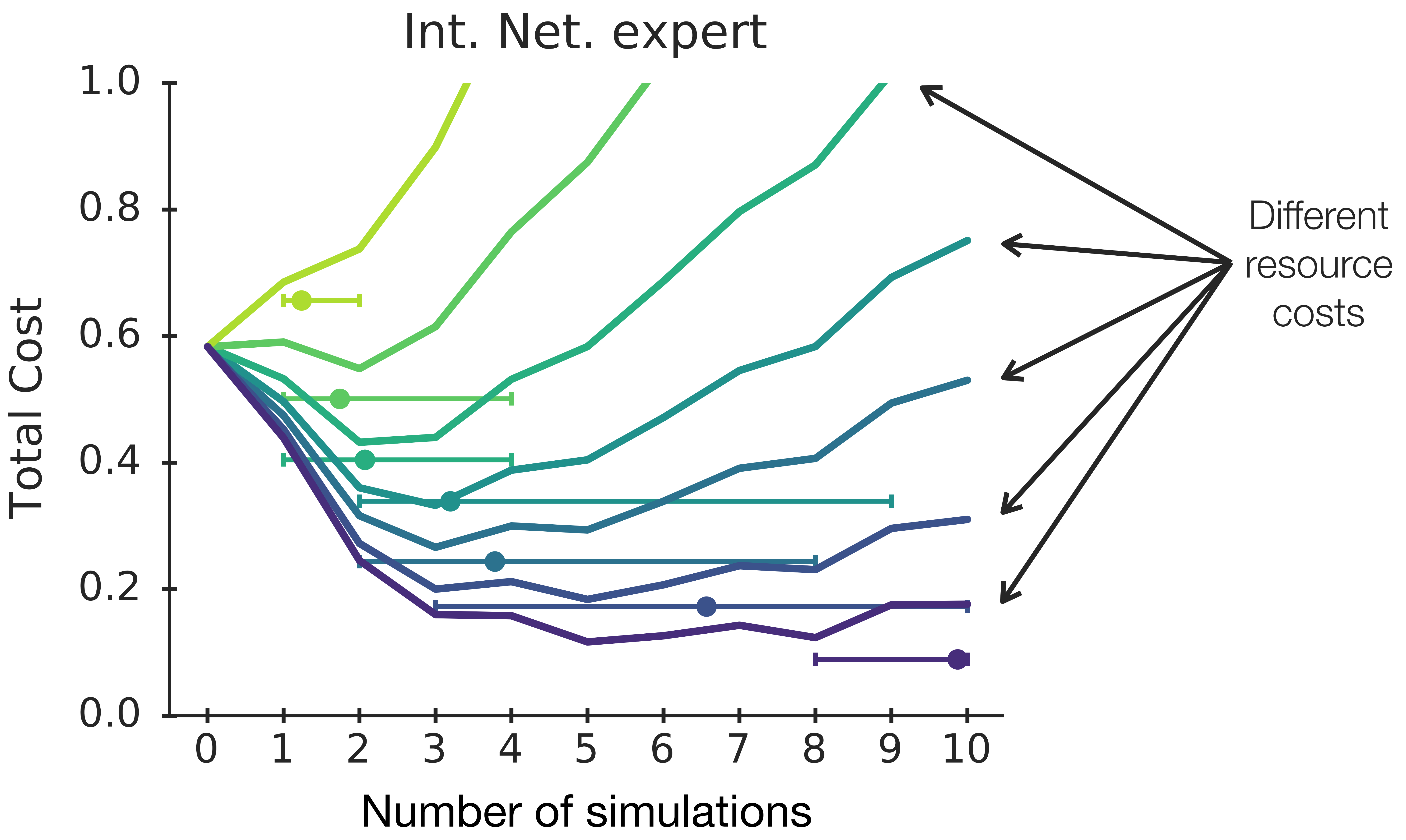
Total cost (five planets)



Different resource costs

Metacontroller Results

Total cost (five planets)

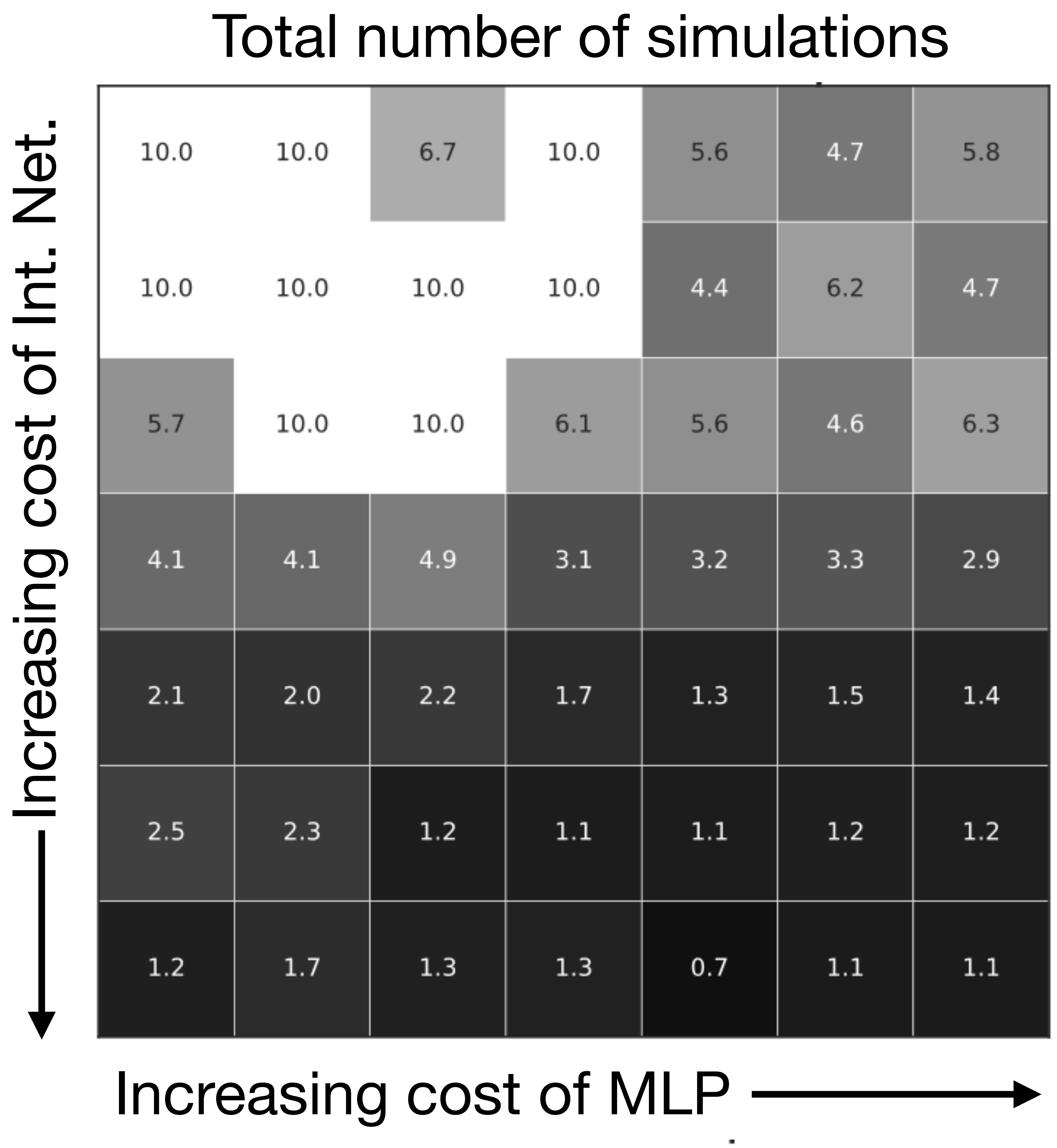


Metacontroller Results

Multiple experts (MLP + IN)

Metacontroller Results

Multiple experts (MLP + IN)



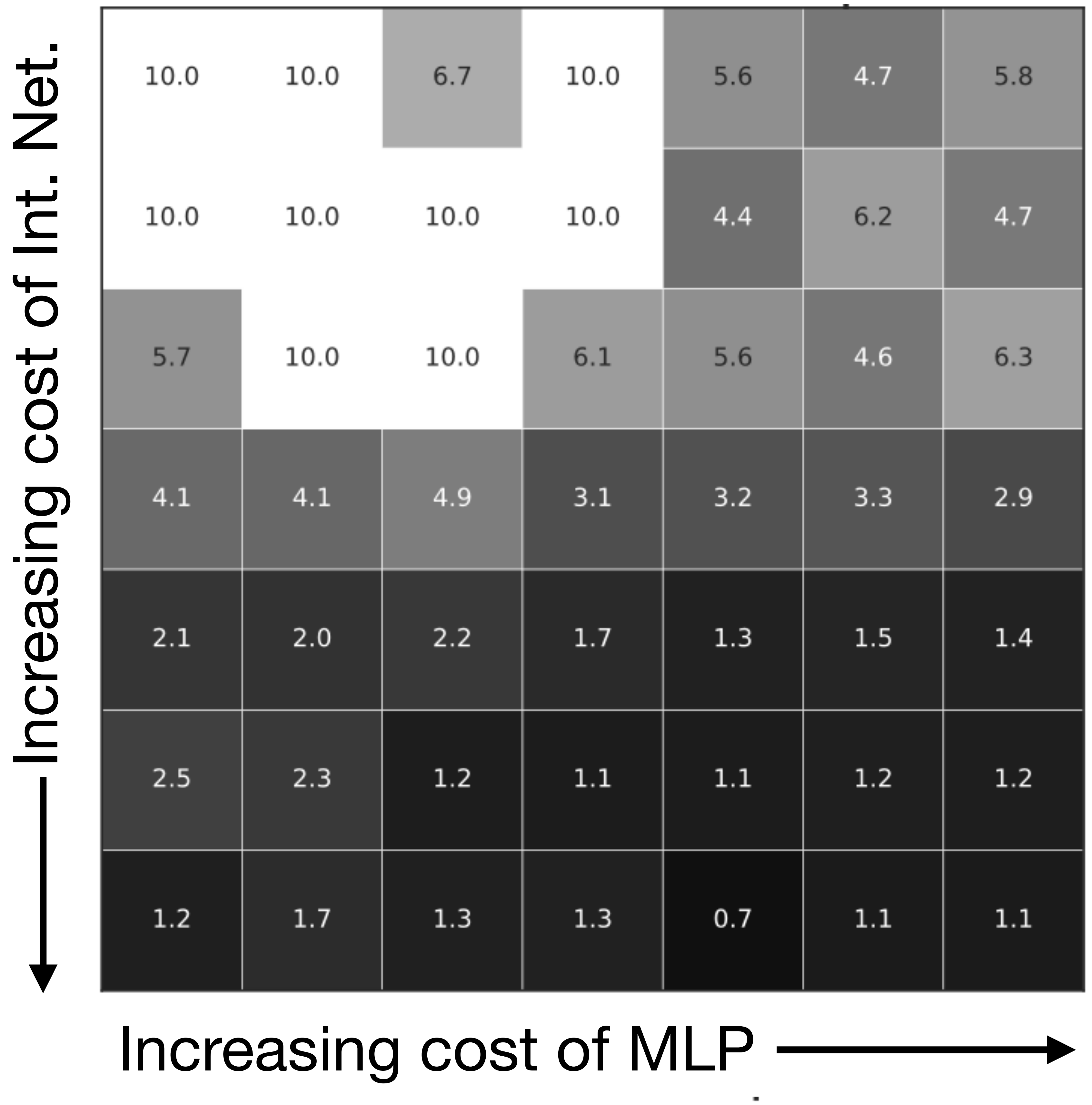
Total number of simulations

10.0	10.0	6.7	10.0	5.6	4.7	5.8
10.0	10.0	10.0	10.0	4.4	6.2	4.7
5.7	10.0	10.0	6.1	5.6	4.6	6.3
4.1	4.1	4.9	3.1	3.2	3.3	2.9
2.1	2.0	2.2	1.7	1.3	1.5	1.4
2.5	2.3	1.2	1.1	1.1	1.2	1.2
1.2	1.7	1.3	1.3	0.7	1.1	1.1

Metacontroller Results

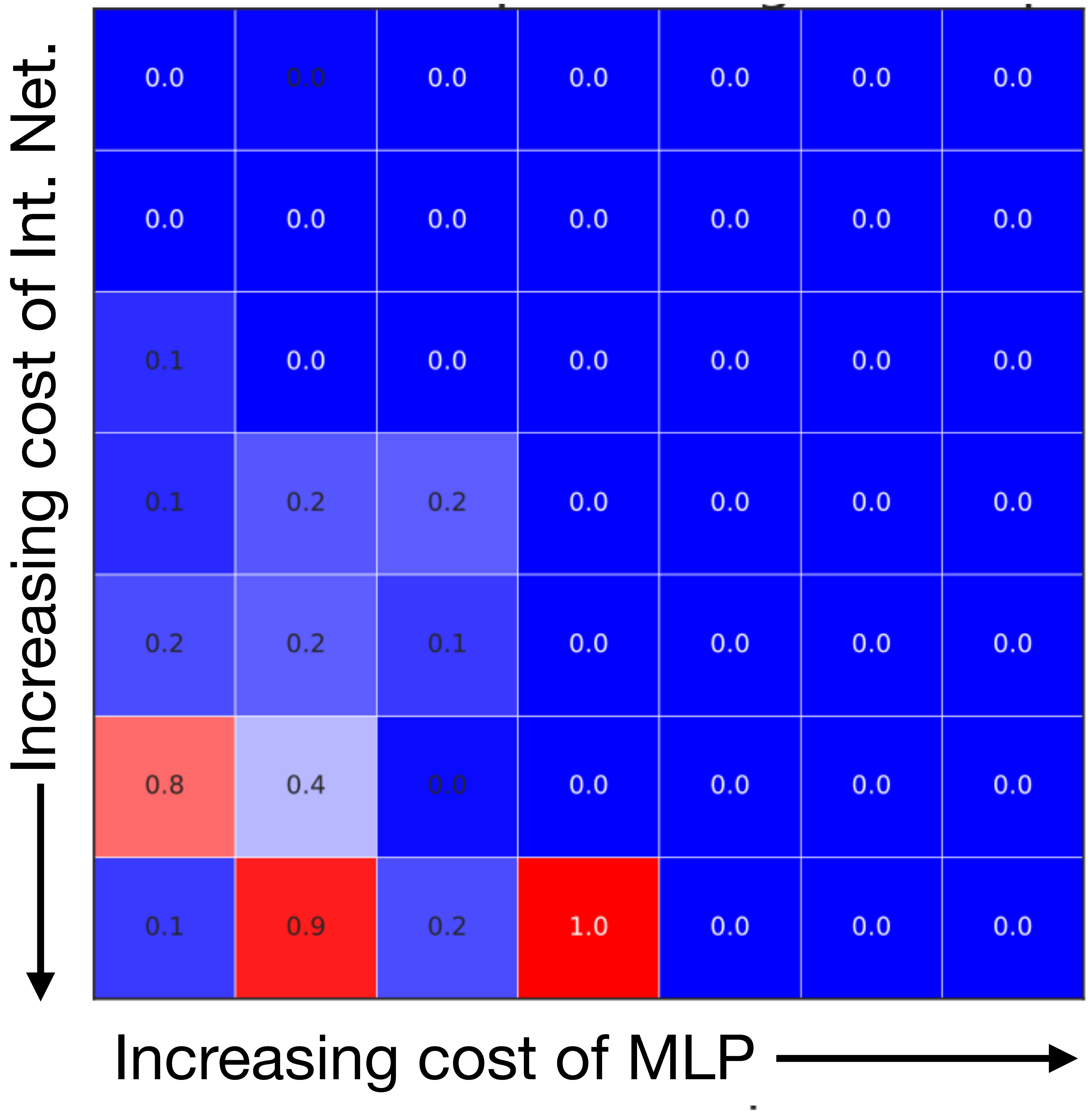
Multiple experts (MLP + IN)

Total number of simulations



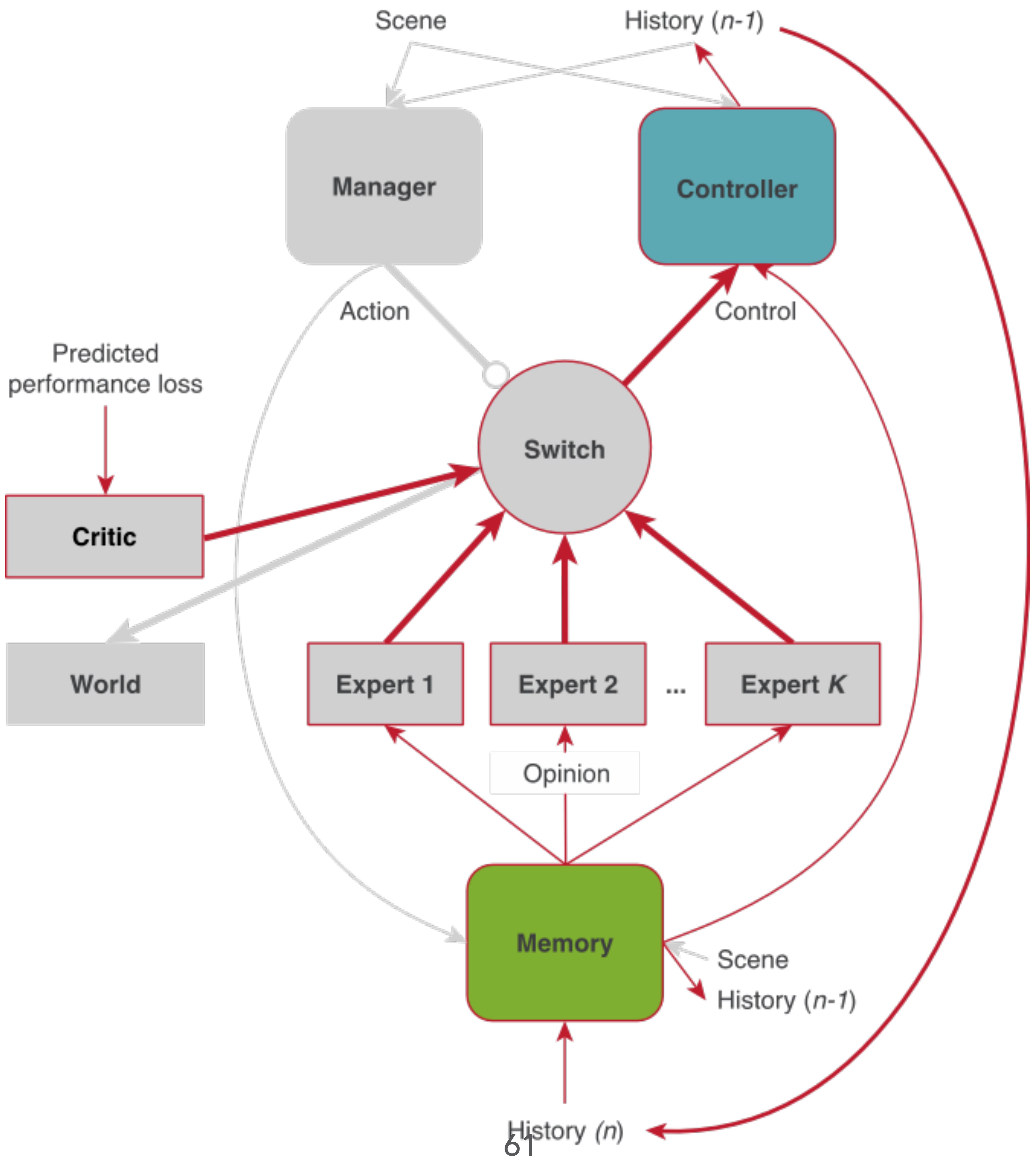
10.0	10.0	6.7	10.0	5.6	4.7	5.8
10.0	10.0	10.0	10.0	4.4	6.2	4.7
5.7	10.0	10.0	6.1	5.6	4.6	6.3
4.1	4.1	4.9	3.1	3.2	3.3	2.9
2.1	2.0	2.2	1.7	1.3	1.5	1.4
2.5	2.3	1.2	1.1	1.1	1.2	1.2
1.2	1.7	1.3	1.3	0.7	1.1	1.1

Fraction of sims using MLP expert

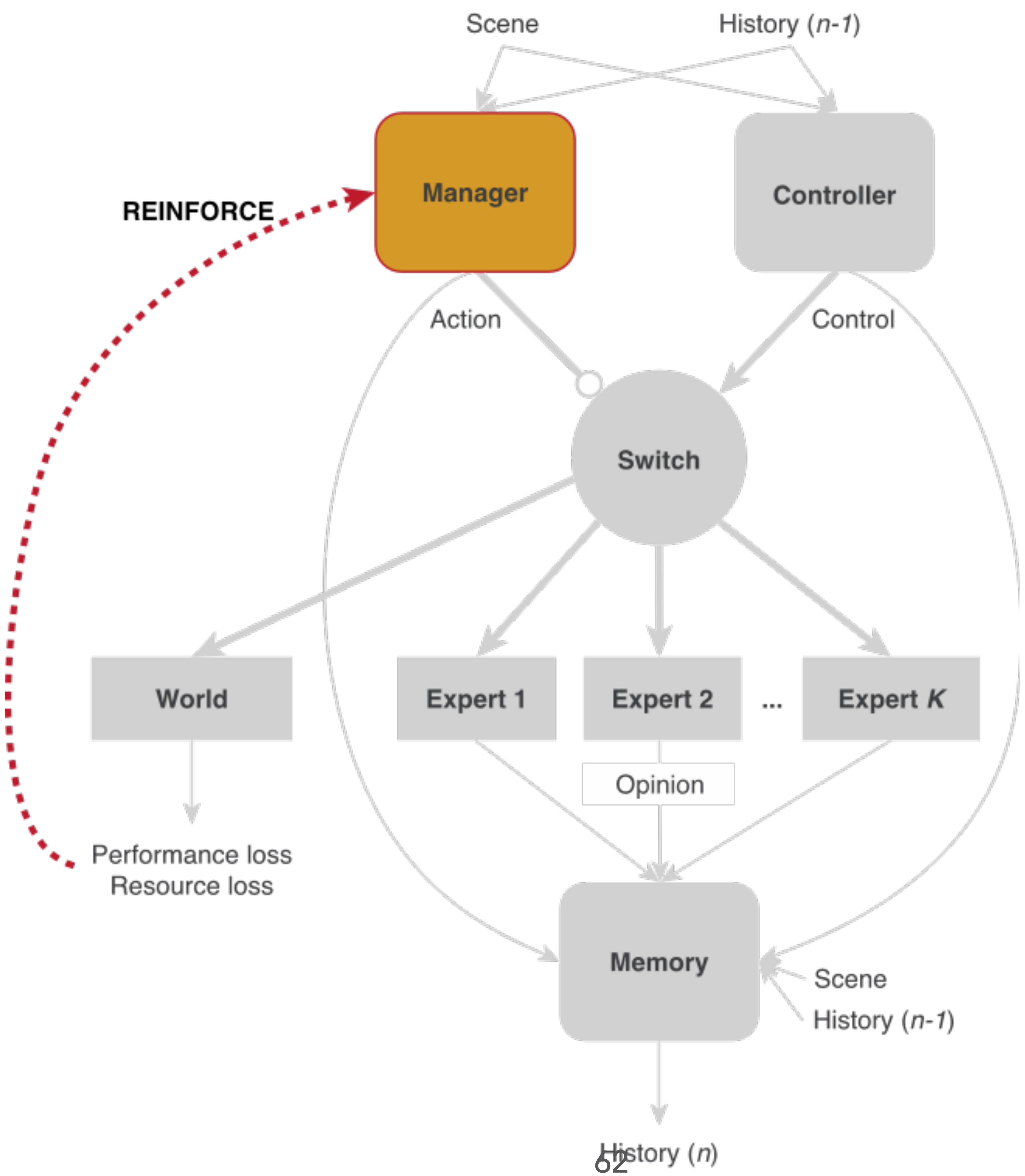


0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.1	0.0	0.0	0.0	0.0	0.0	0.0
0.1	0.2	0.2	0.0	0.0	0.0	0.0
0.2	0.2	0.1	0.0	0.0	0.0	0.0
0.8	0.4	0.0	0.0	0.0	0.0	0.0
0.1	0.9	0.2	1.0	0.0	0.0	0.0

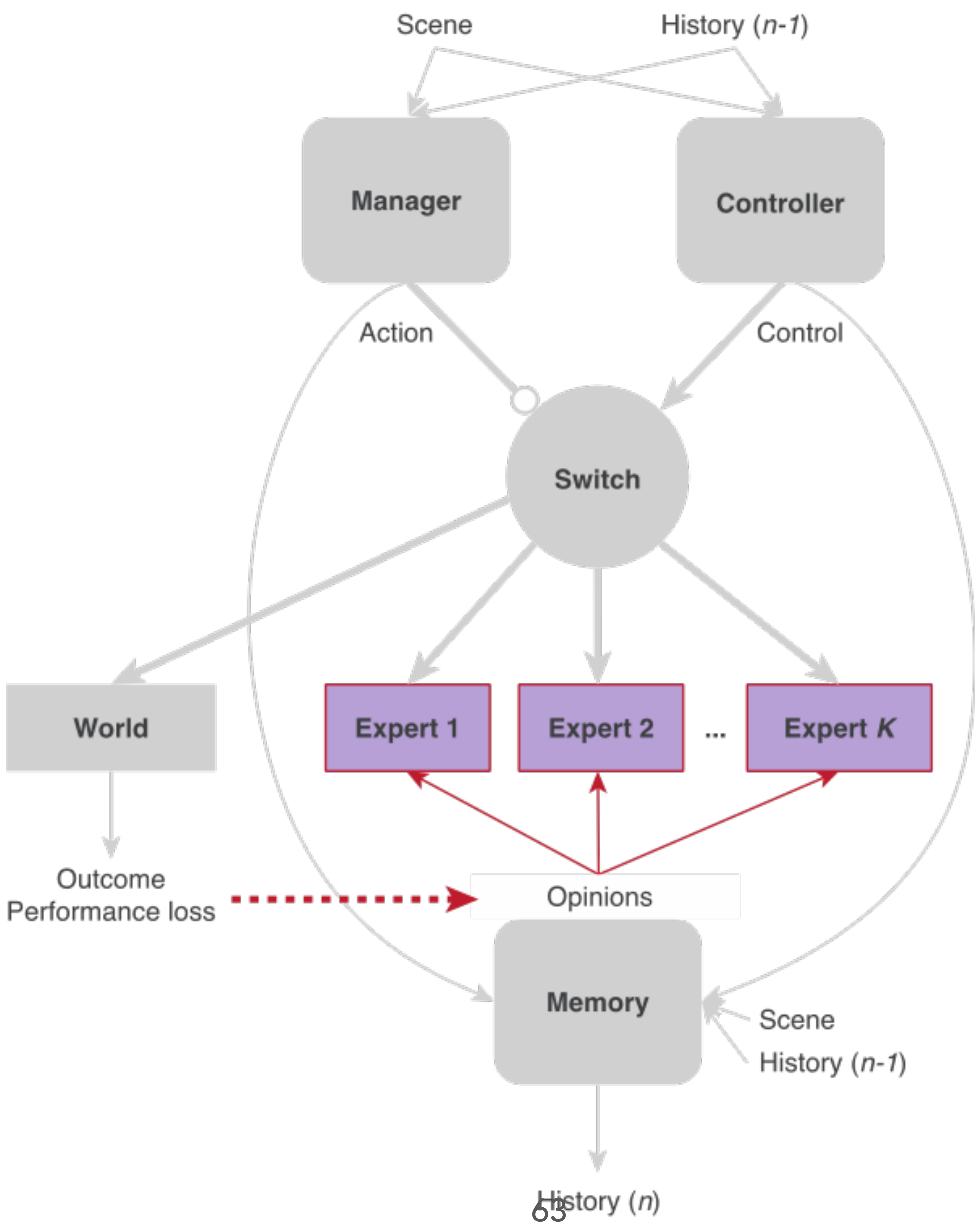
Training the Controller and Memory



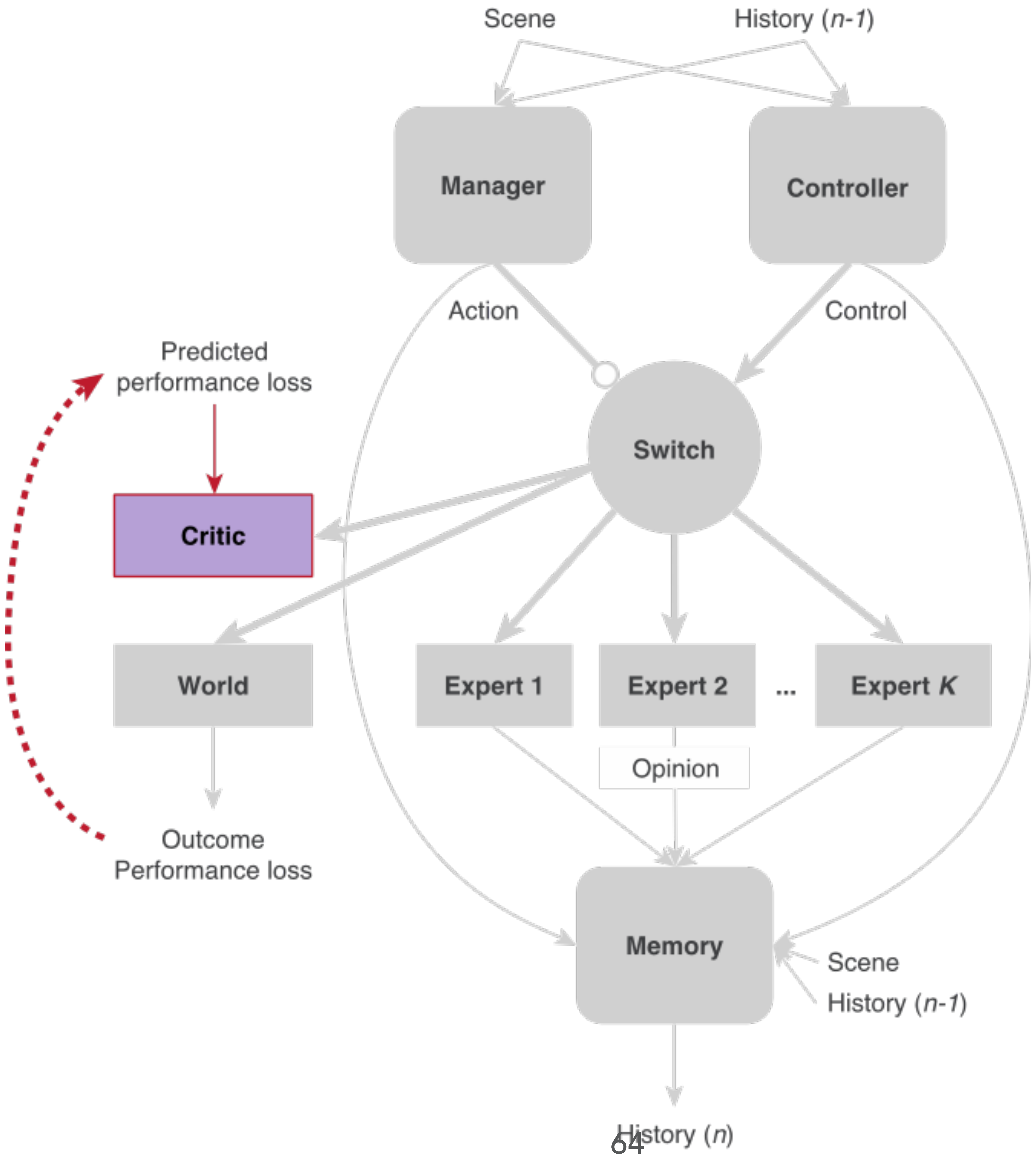
Training the Manager



Training the Experts



Training the Critic



DeepMind

Jessica Hamrick (@jhamrick)