Abstract

Meta-learning, or learning to learn, has gained renewed interest in recent years within the artificial intelligence community. However, meta-learning is incredibly prevalent within nature, has deep roots in cognitive science and psychology, and is currently studied in various forms within neuroscience. In this talk, I will discuss recent work casting previous neuroscientific findings within a meta-learning perspective, as well as the ability of deep learning systems trained through meta-RL to perform more complex forms of cognition, such as causal decision-making.

Bio

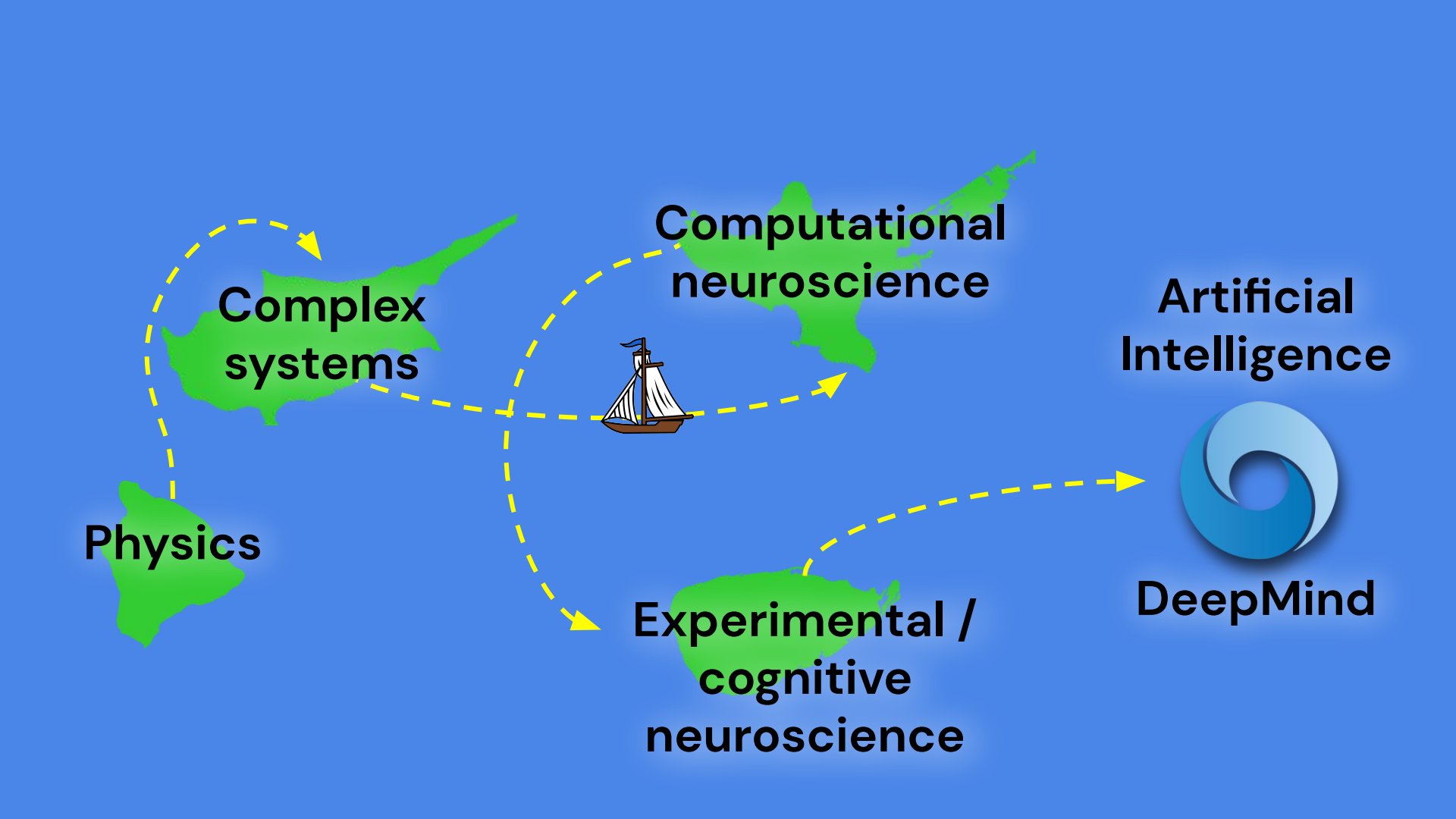
Jane Wang is a senior research scientist at DeepMind on the neuroscience team, working on meta-reinforcement learning and neuroscience-inspired artificial agents. She obtained a Ph.D from the University of Michigan in Applied Physics, where she worked on computational neuroscience models of memory consolidation and complex dynamical systems, and completed a post-doc at Northwestern University, working on cognitive neuroscience of learning and memory systems in humans.

Meta-learning in natural and artificial intelligence

CS330 Guest lecture

Jane X. Wang
November 9, 2020

DeepMind


Complex systems
Computational neuroscience
Artificial Intelligence
Physics
DeepMind
Experimental / cognitive neuroscience

What I hope to convince you of

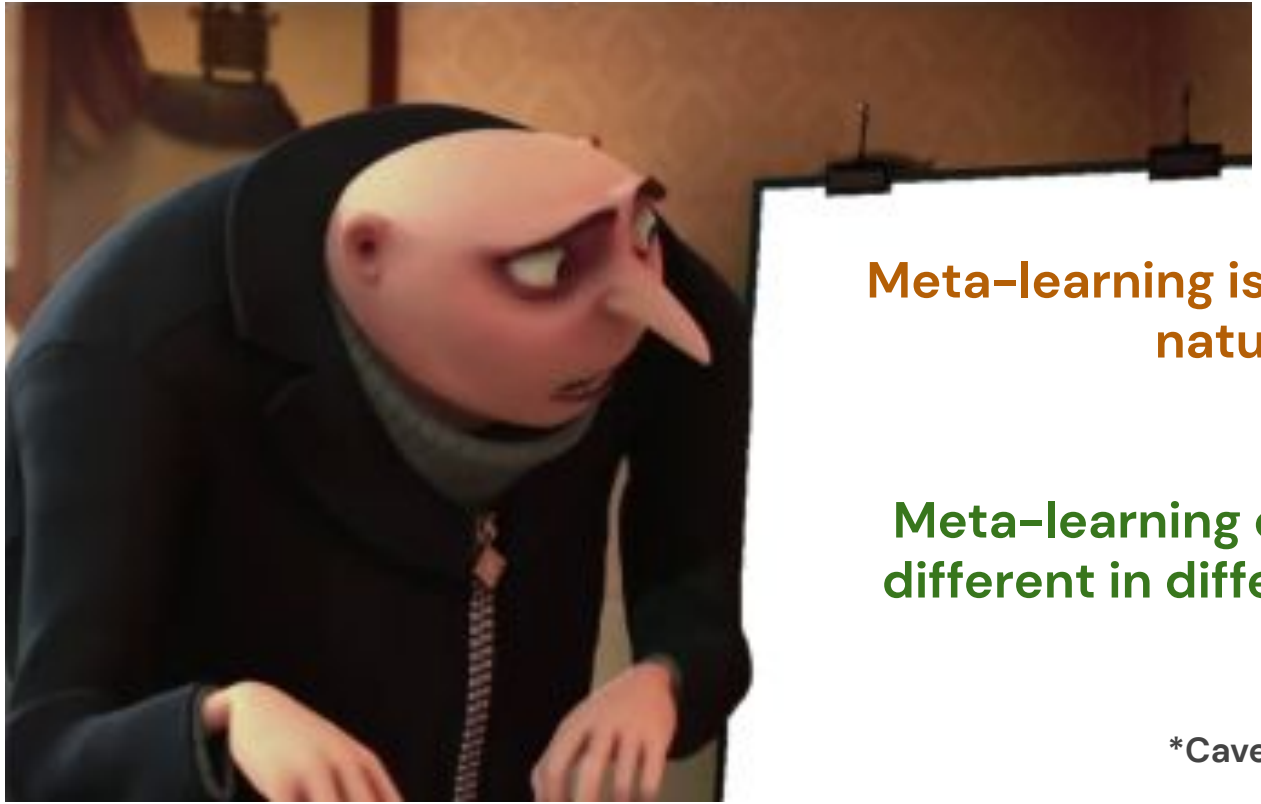
Meta-learning is the default in nature

What I hope to convince you of

Meta-learning is the default in nature

Meta-learning can look very different in different settings

What I hope to convince you of



Meta-learning is the default in nature

Meta-learning can look very different in different settings

*Caveat

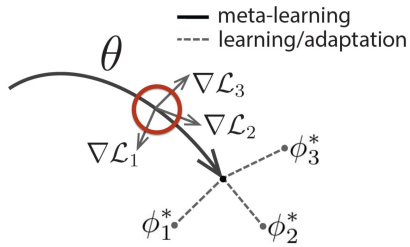
What meta-learning looks like in ML

Optimization-based

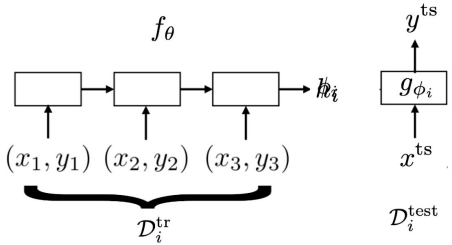
$$\min_{\theta} \sum_{\text{task } i} \mathcal{L}(\theta - \alpha \nabla_\theta \mathcal{L}(\theta, \mathcal{D}_i^{\text{tr}}), \mathcal{D}_i^{\text{ts}})$$

θ parameter vector being meta-learned

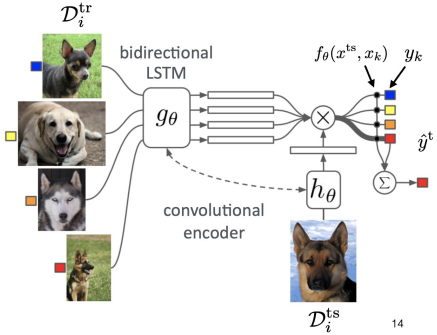
ϕ_i^* optimal parameter vector for task i



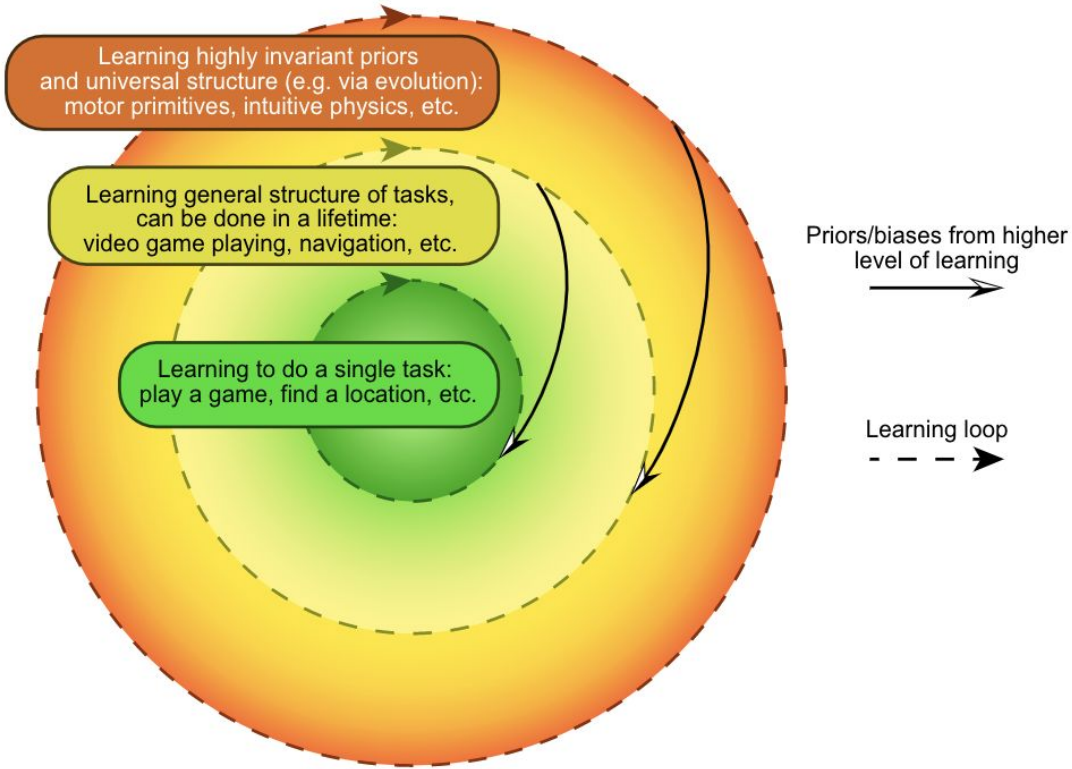
Blackbox (LSTM)



Nonparametric



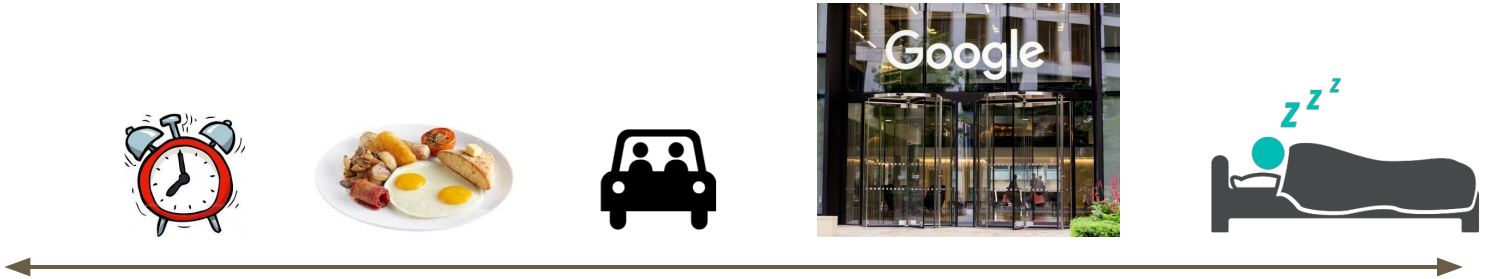
Multiple nested timescales of learning in nature



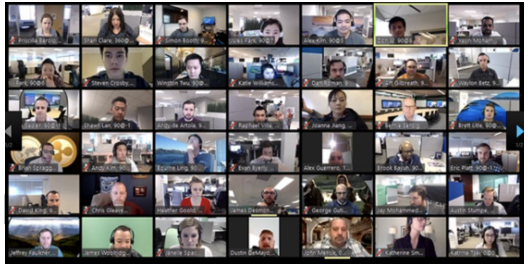
What does meta-learning look like in nature?

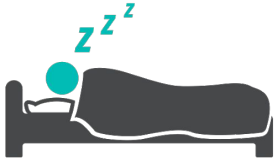
Priors learned from previous experience helps
to inform faster learning / better decisions

What does meta-learning look like in one day?

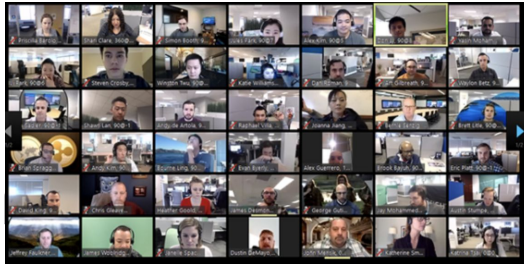


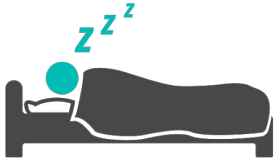
What does meta-learning look like in one day?





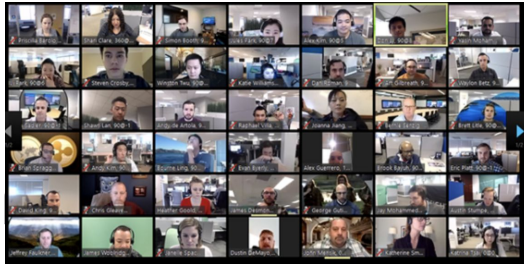
What does meta-learning look like in one day?

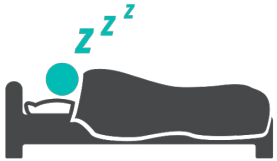




Espresso	5.90
Iced Coffee	4.95
Macchiato	5.95
Coffee with Cream	5.95
Chocolate Coffee	4.25
Long Black	2.99
House Blend	2.99
Decaf Coffee	2.99
Latte	4.95

What does meta-learning look like in one day?



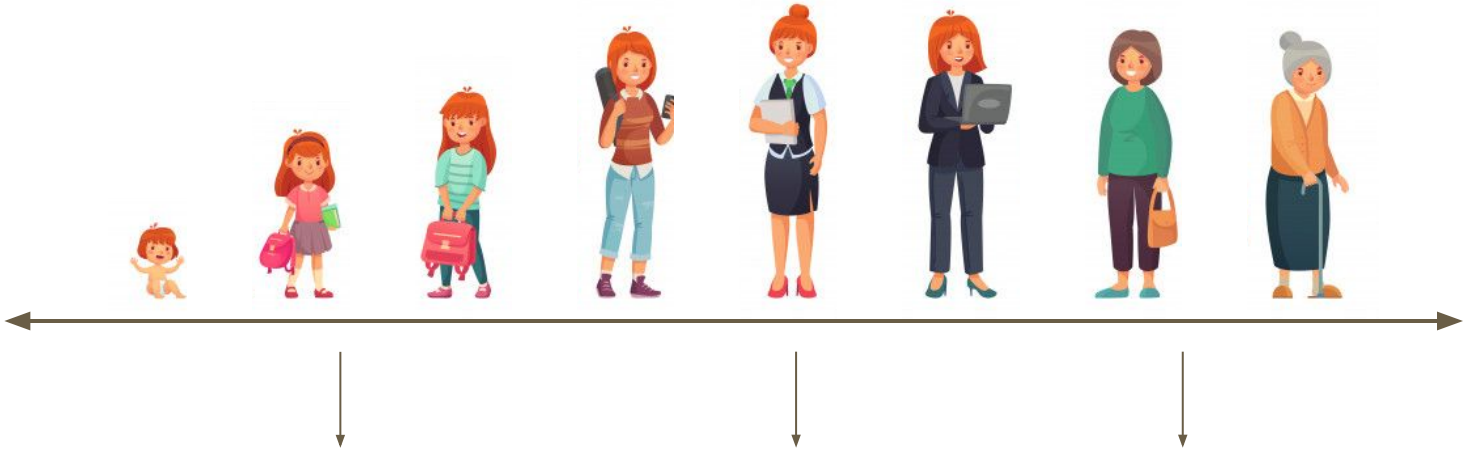


Espresso	5.90
Iced Coffee	4.95
Macchiato	5.95
Coffee with Cream	5.95
Chocolate Coffee	4.25
Long Black	2.99
House Blend	2.99
Decaf Coffee	2.99
Latte	4.95

Prior = Coffee shops tend to be consistent in quality

Learned decision = come back tomorrow

What does meta-learning look like in one lifetime?

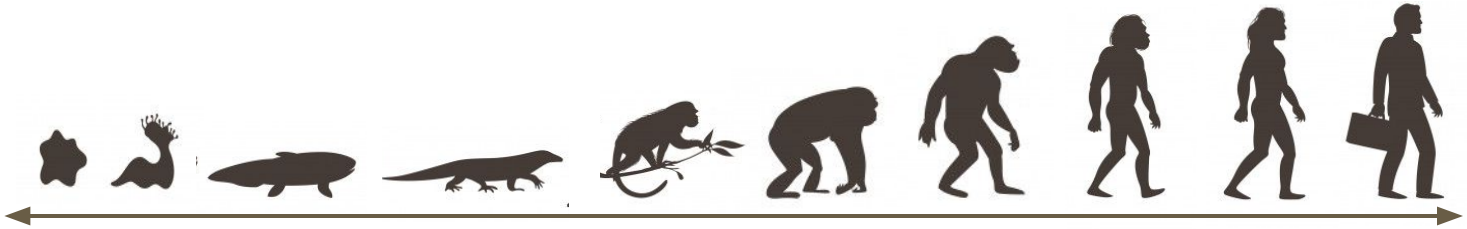


Learning: Language, social skills, motor skills | Knowledge, career choice | Lifelong skills

Priors = Propensity for language, intuitive physics, motor primitives, biological wiring

Image: freepik.com

What does meta-learning look like in one (evolutionary) epoch?

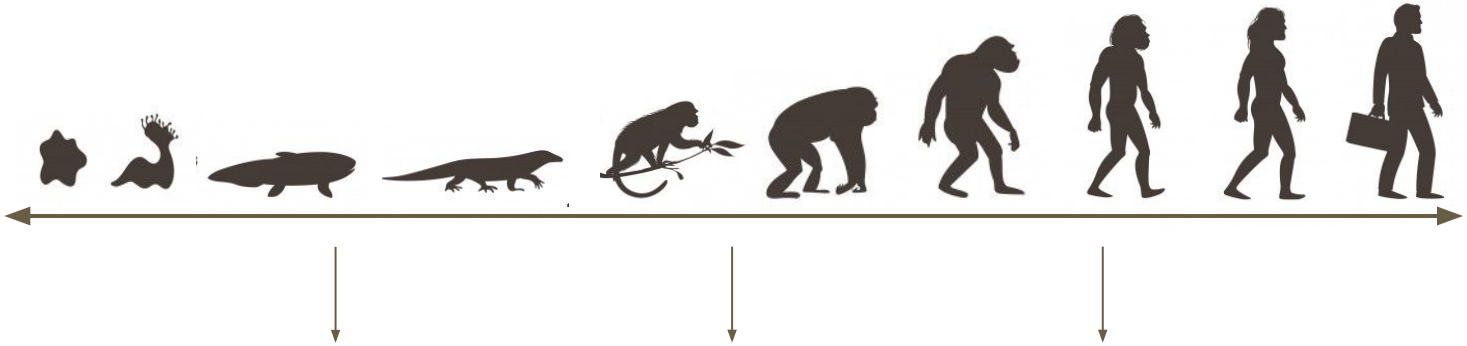


Learning: Survival adaptation | Developmental trajectories | Intuitive physics

Priors = ?

Image: freepik.com

What does meta-learning look like in one (evolutionary) epoch?



Perspective | Open Access | Published: 21 August 2019

A critique of pure learning and what artificial neural networks can learn from animal brains

Anthony M. Zador

Nature Communications **10**, Article number: 3770 (2019) | Cite this article

Image: freepik.com

A spectrum of fast and slow learning in biological organisms

Fast to mature

Purely innate behavior

Small range of behaviors

Slow to mature

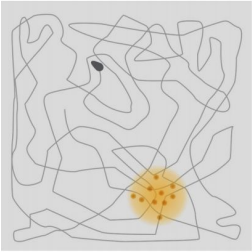
Learned + innate behavior

Large range of behaviors

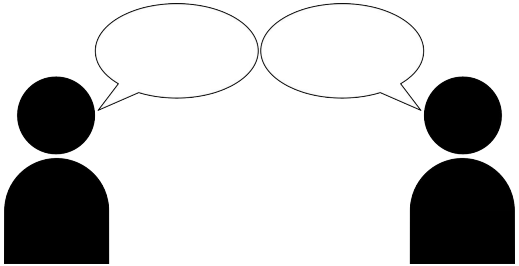
Two types of learning we can study in neuroscience

1. **Innate behaviors - prespecified from birth**

Place cells



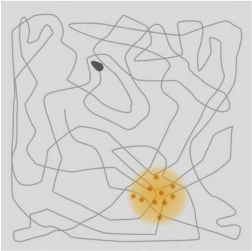
nobelprize.org



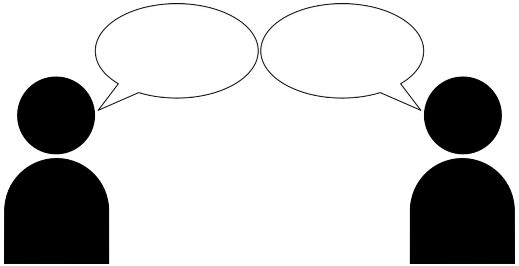
Two types of learning we can study in neuroscience

1. **Innate behaviors - prespecified from birth**

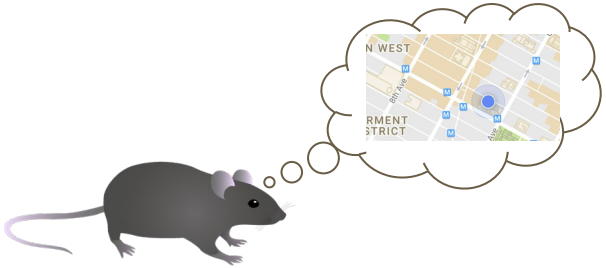
Place cells



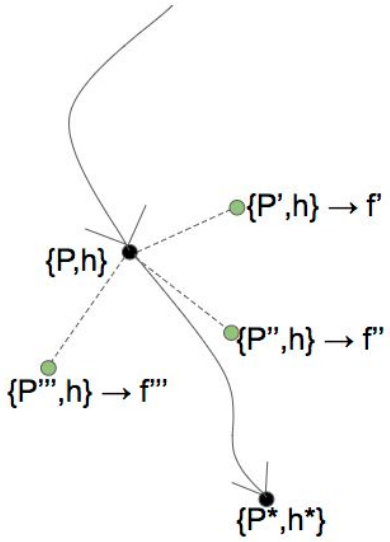
nobelprize.org



2. **Learned behaviors - fast adaptation (ie specific place fields, item-context association), can arise out of innate processes**



The Baldwin effect



Baldwinian Evolution

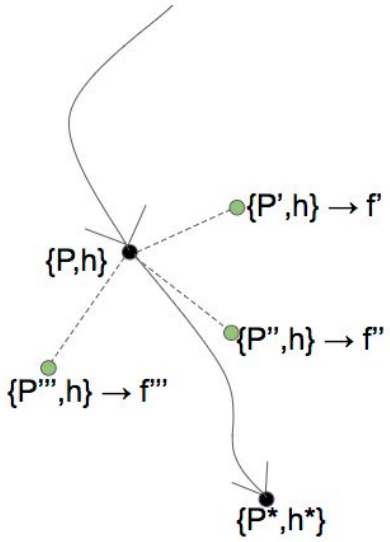
Meta-learning by the Baldwin Effect, Fernando et al, 2018 GECCO

"If animals entered a new environment—or their old environment rapidly changed—those that could flexibly respond by learning new behaviors or by ontogenetically adapting would be naturally preserved. This saved remnant would, over several generations, have the opportunity to exhibit spontaneously congenital variations similar to their acquired traits and have these variations naturally selected."

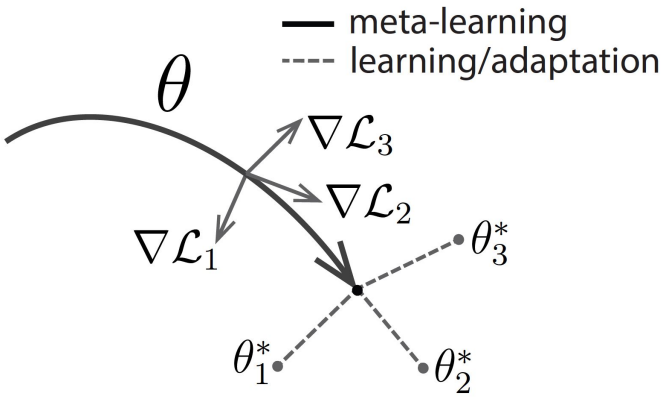
Darwin and the Emergence of Evolutionary Theories of Mind and Behavior. Richards, Robert J. (1987).

A new factor in evolution, J Mark Baldwin. (1896).

How learning can guide evolution. Hinton, Geoffrey E.; Nowlan, Steven J. (1987). Complex Systems. 1: 495–502.



Baldwinian Evolution



Learn the initial parameters of a neural network such that, within just a few steps of gradient descent (weight adjustment), you can solve a variety of new tasks

Meta-learning by the Baldwin Effect,
Fernando et al, 2018 GECCO

Model-agnostic meta-learning
Finn et al, 2017 ICML

What I hope to convince you of

Meta-learning is the default in nature

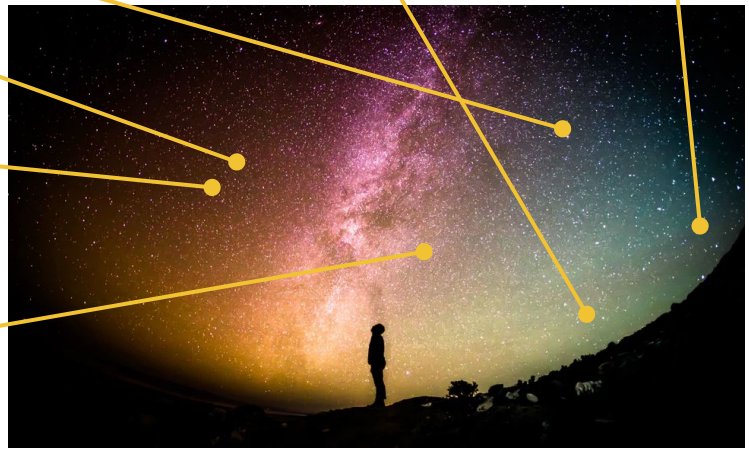
What I hope to convince you of

Meta-learning is the default in nature

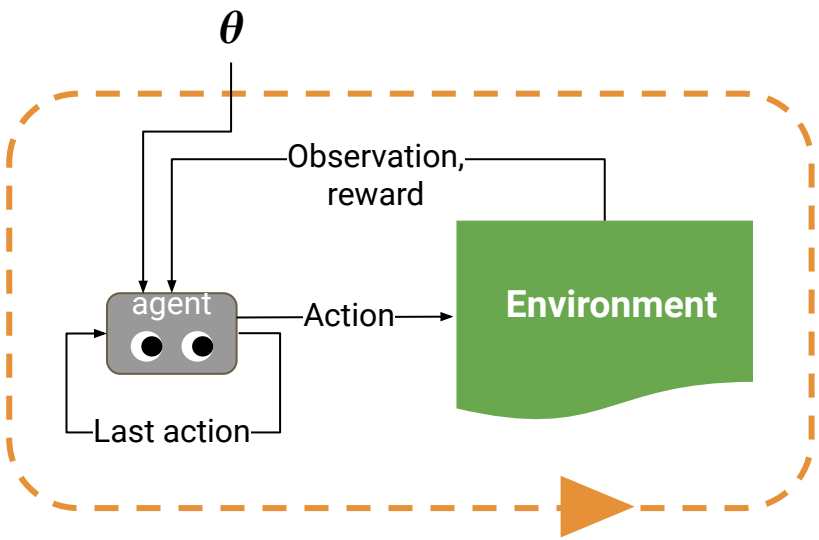
Meta-learning can look very different in different settings

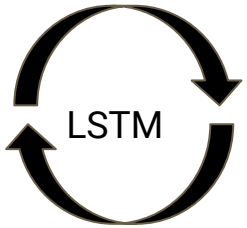
It’s all in the task distribution

A structured universe of tasks = structured priors

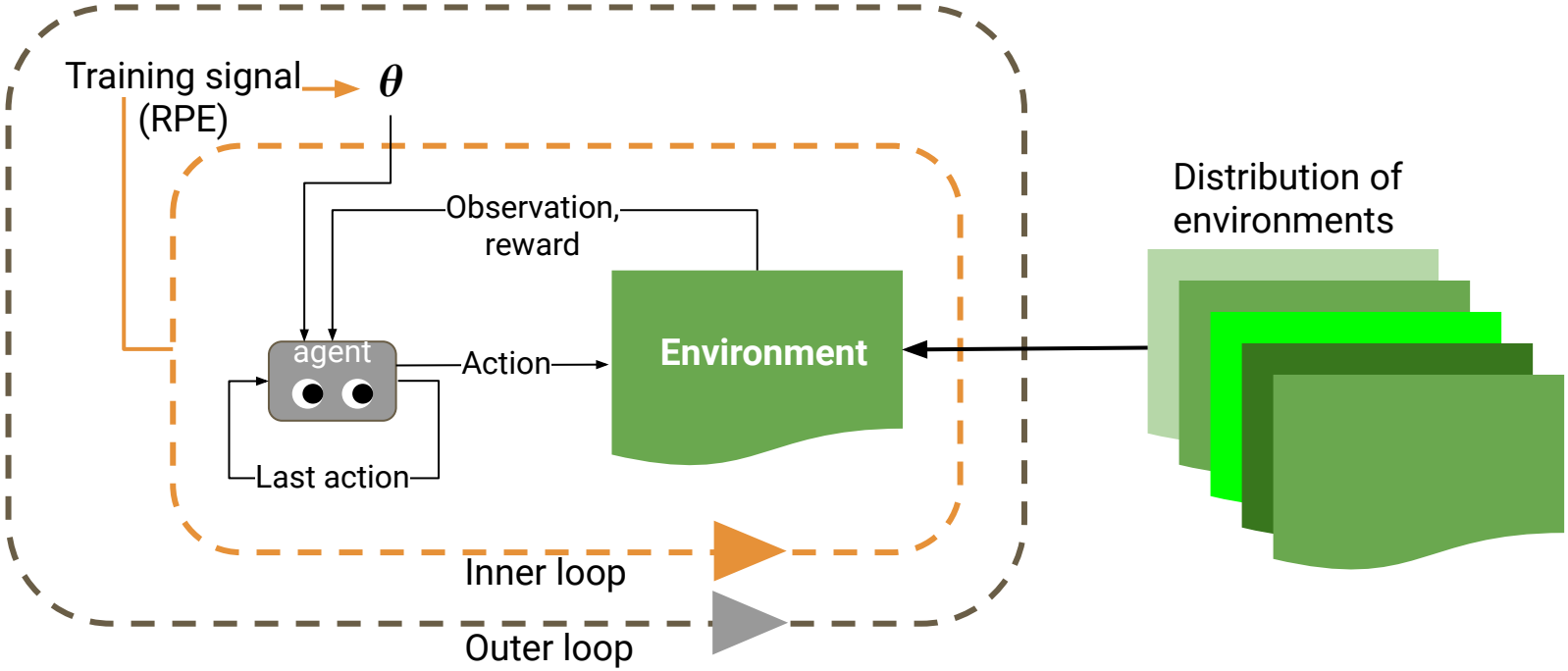


Memory-based learning to reinforcement learn (L2RL)

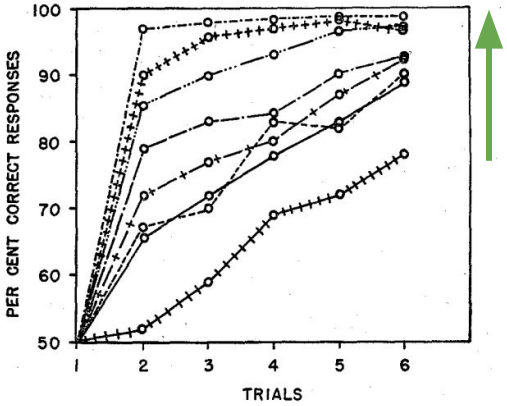




Memory-based learning to reinforcement learn (L2RL)



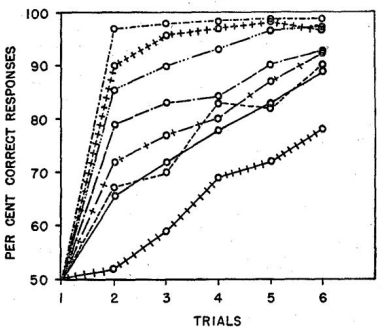
The “Harlow task”



Training episodes

*Harlow, **1949(!)**, Psychological Review*

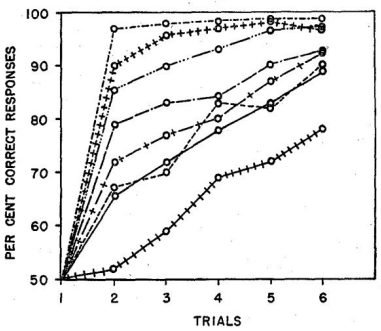
Animal

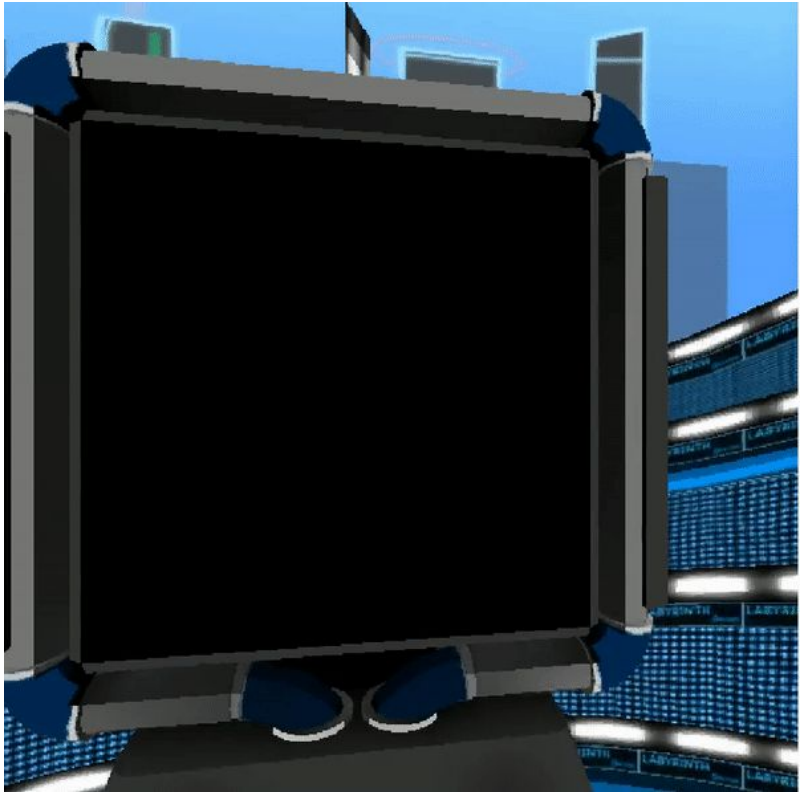


Wang et al. Nature Neuroscience (2018)

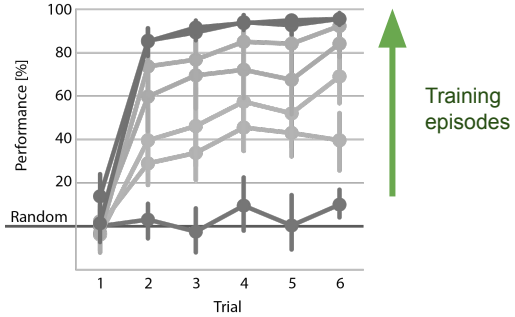
Behavior with weights of NN frozen

Animal



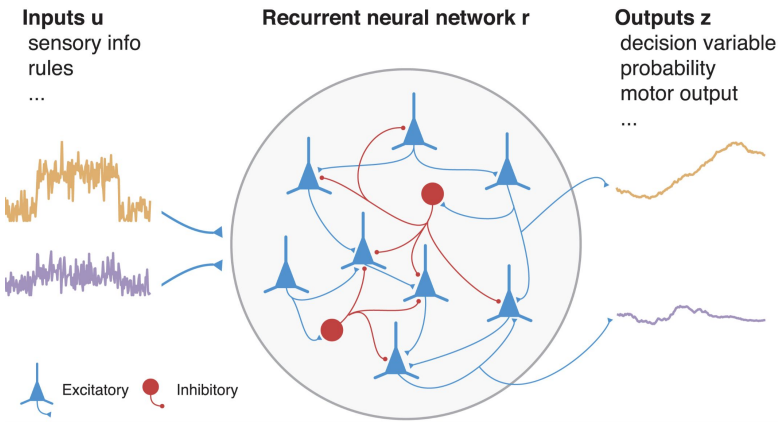


Artificial agent



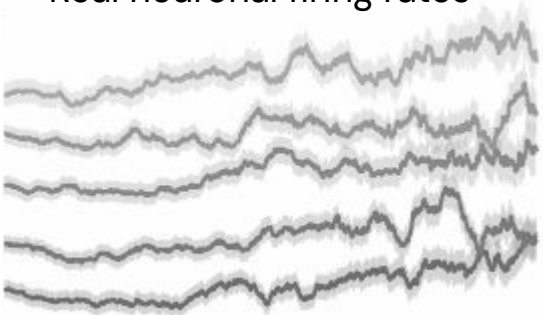
Wang et al. Nature Neuroscience (2018)

Memory-based meta-learning implements the inner loop of learning via the hidden states of the recurrent neural network, providing a nice correspondence with neural activations



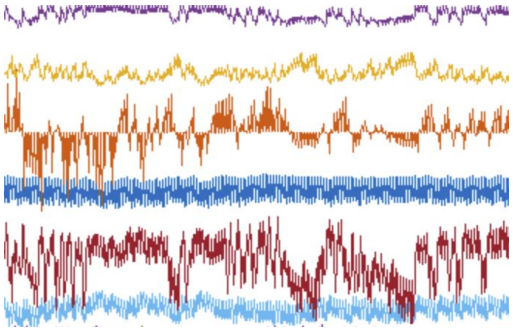
Song et al. PLoS Comput Biol (2016)

Real neuronal firing rates



Bari et al. Neuron (2019)

LSTM hidden states



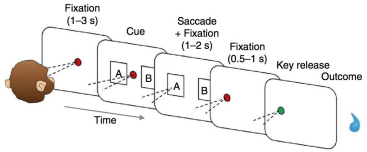
Memory-based meta-learning captures real behavior and neural dynamics

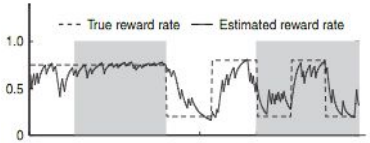
Articles
https://doi.org/10.1038/s41593-018-0147-8
nature neuroscience

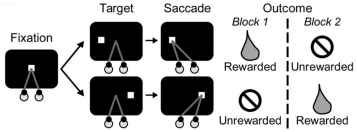
Prefrontal cortex as a meta-reinforcement learning system

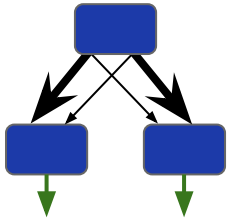
Jane X. Wang[1,5], Zeb Kurth-Nelson[1,2,5], Dharshan Kumaran[1,3], Dhruva Tirumala[1], Hubert Soyer[1], Joel Z. Leibo[1], Demis Hassabis[1,4] and Matthew Botvinick[1,4*]

Over the past 20 years, neuroscience research on reward-based learning has converged on a canonical model, under which the neurotransmitter dopamine 'stamps in' associations between situations, actions and rewards by modulating the strength of synaptic connections between neurons. However, a growing number of recent findings have placed this standard model under strain. We now draw on recent advances in artificial intelligence to introduce a new theory of reward-based learning. Here, the

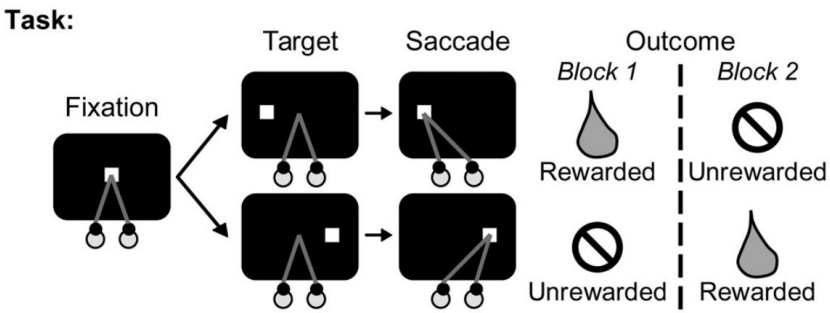






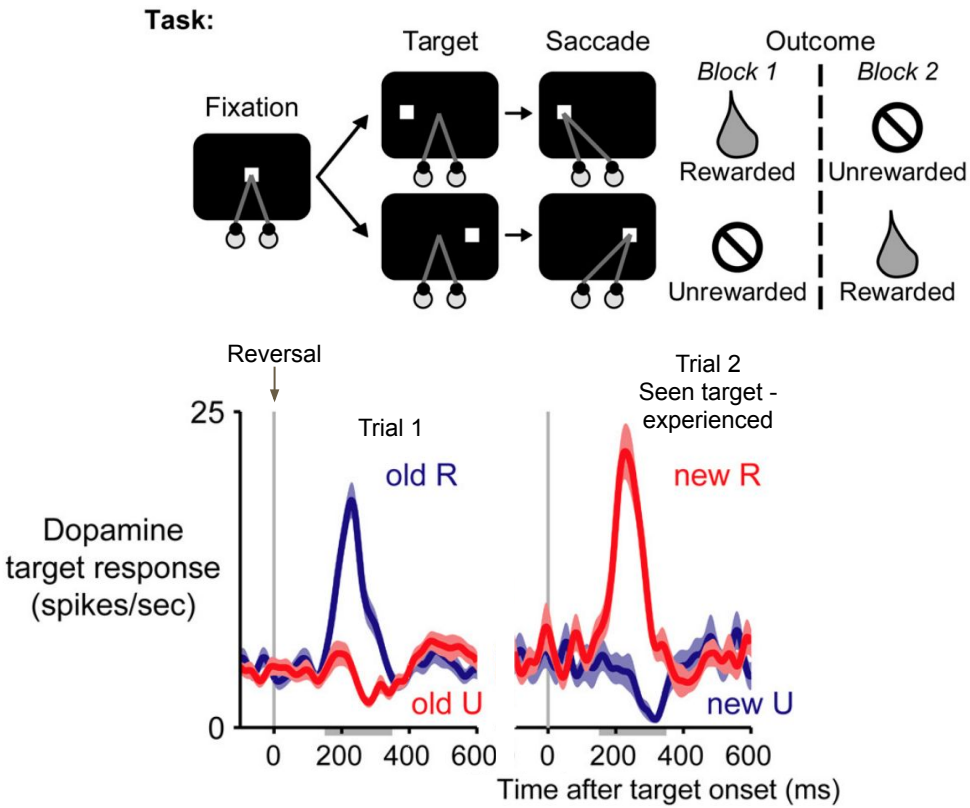


Dopamine reward prediction errors (RPEs) reflect indirect, inferred value



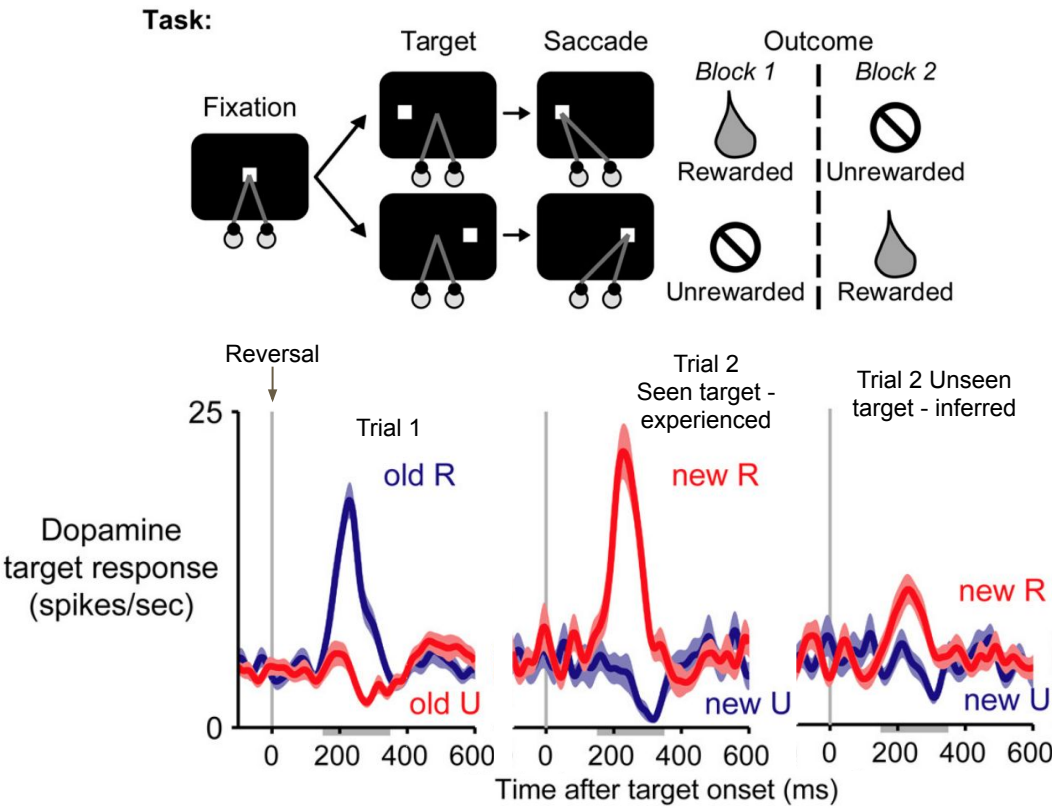
Bromberg-Martin et al., J Neurophys, 2010

Dopamine reward prediction errors (RPEs) reflect indirect, inferred value



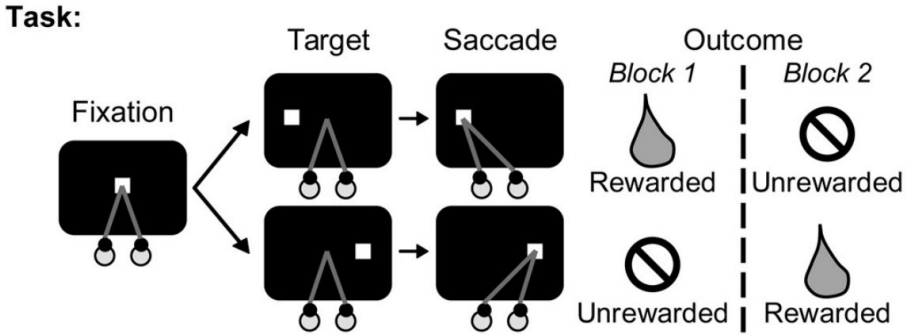
Bromberg-Martin et al., J Neurophys, 2010

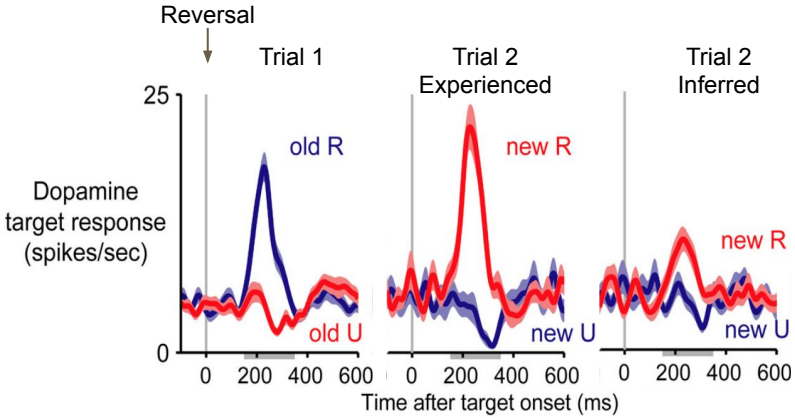
Dopamine reward prediction errors (RPEs) reflect indirect, inferred value



Bromberg-Martin et al., J Neurophys, 2010

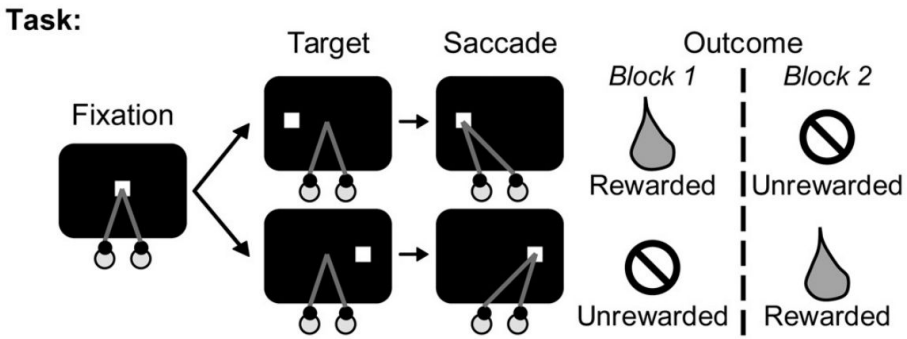
Reward prediction error signal reflects model-based inference



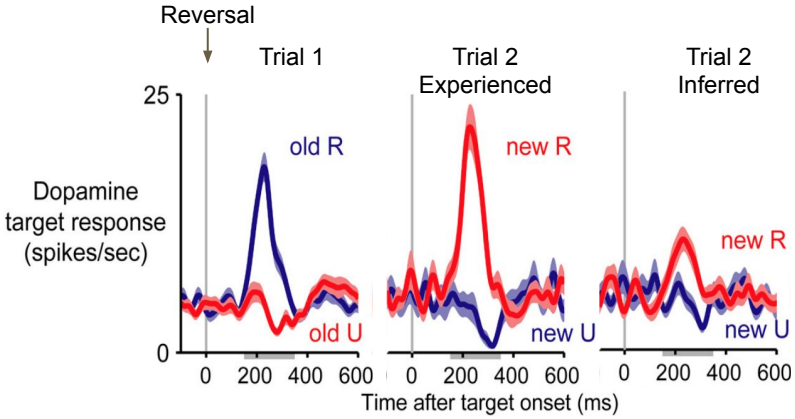


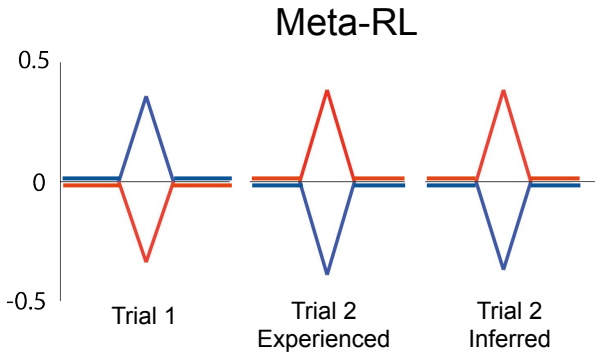
Bromberg-Martin et al, J Neurophys, 2010

Reward prediction error signal reflects model-based inference

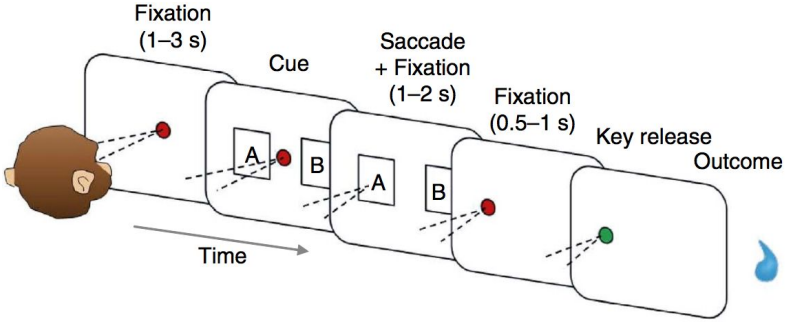


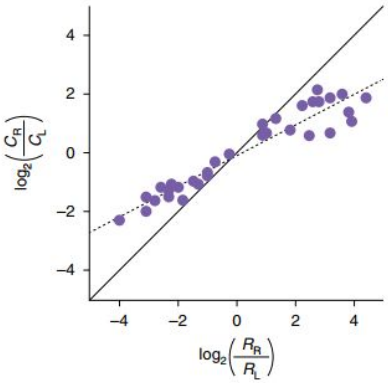
Bromberg-Martin et al, J Neurophys, 2010





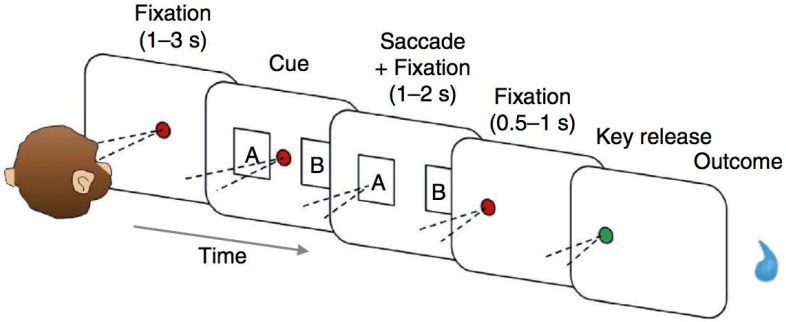
PFC activity dynamics encode information to perform RL

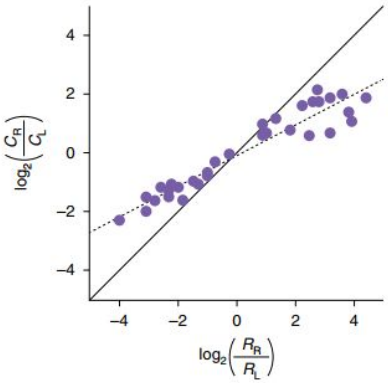




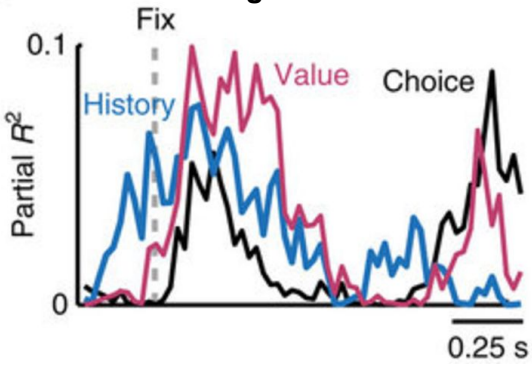
Tsutsui, Grabenhorst, Kobayashi & Schultz, Nature Communications, 2016

PFC activity dynamics encode information to perform RL

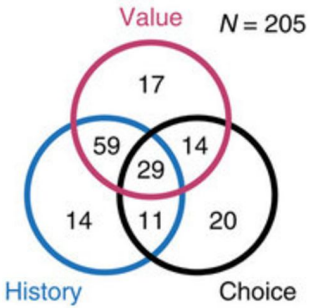




Single neuron

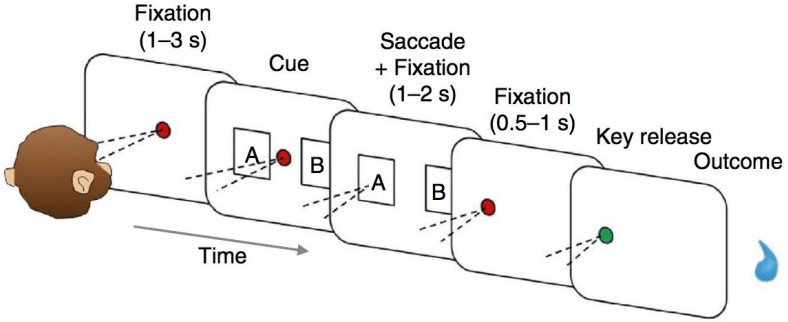


 # Neurons coding for variable

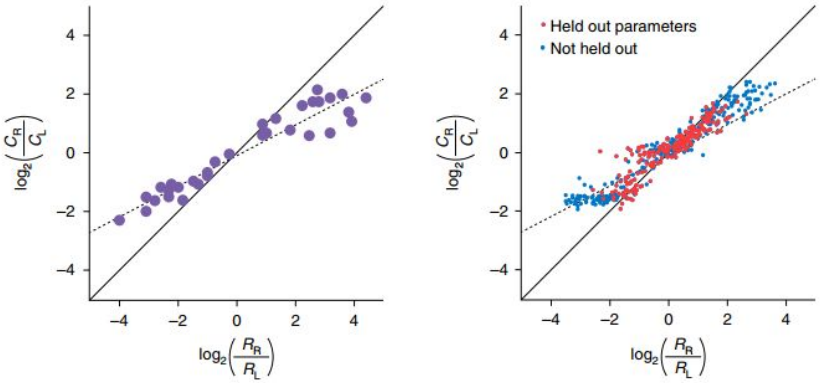


Tsutsui, Grabenhorst, Kobayashi & Schultz, Nature Communications, 2016

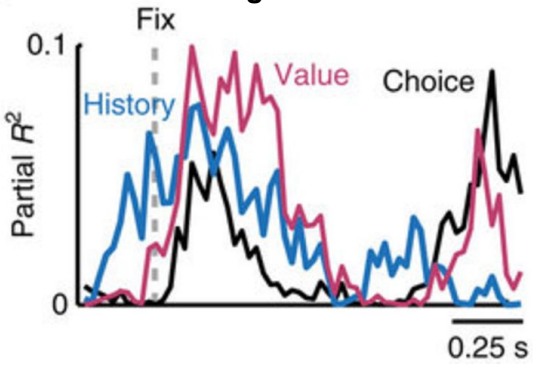
PFC activity dynamics encode information to perform RL



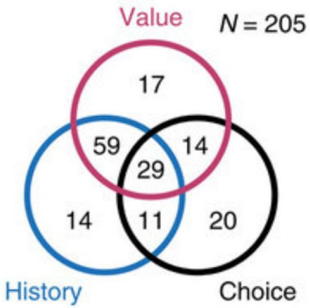
Meta-RL



Single neuron

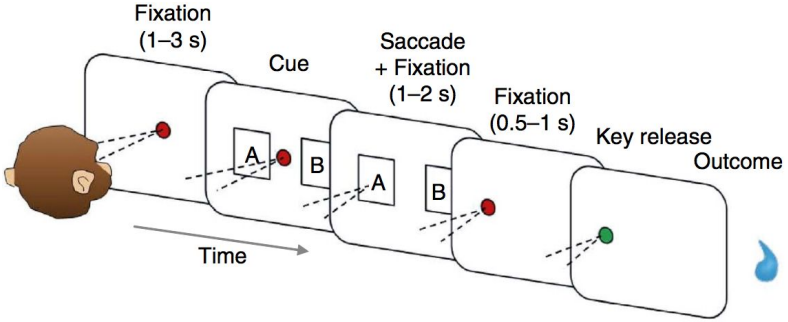


 # Neurons coding for variable

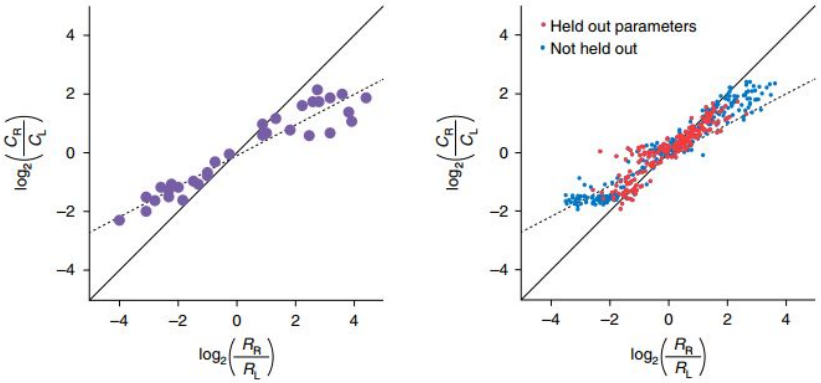


Wang et al. Nature Neuroscience, 2018

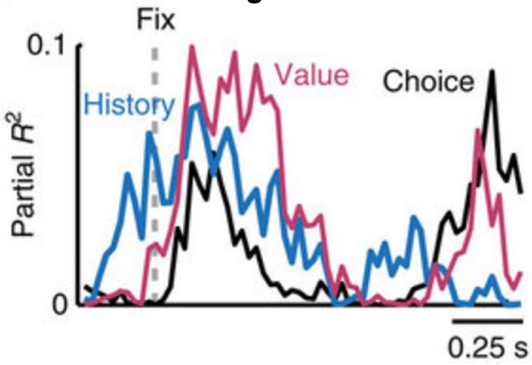
PFC activity dynamics encode information to perform RL



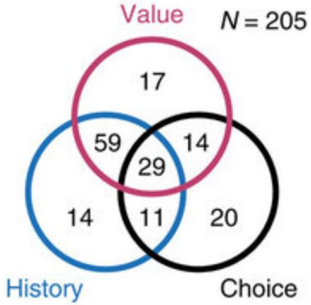
Meta-RL



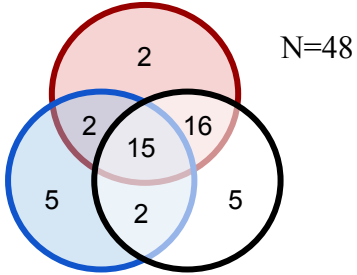
Single neuron



 # Neurons coding for variable



Meta-RL

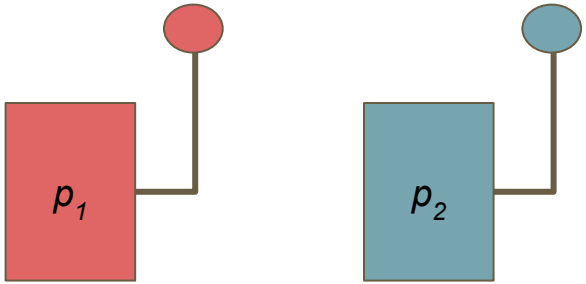


Wang et al. Nature Neuroscience, 2018

2-armed bandits

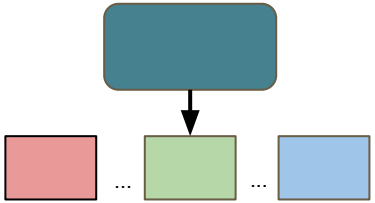
2-armed bandits
independently drawn from
uniform Bernoulli distribution

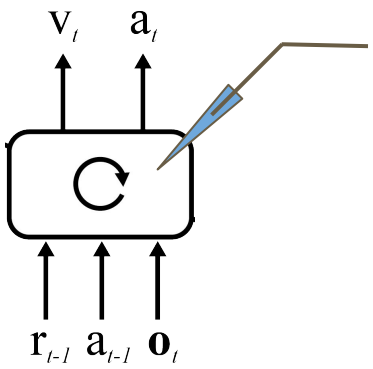
Held constant for 100 trials
=1 episode



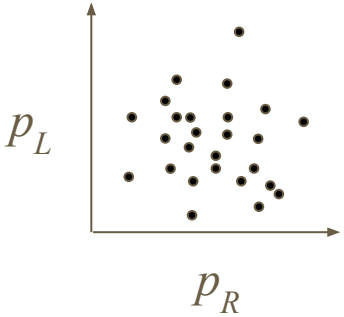
p_i = probability of payout,
drawn uniformly from [0,1],

Agent's neural network internalizes task structure

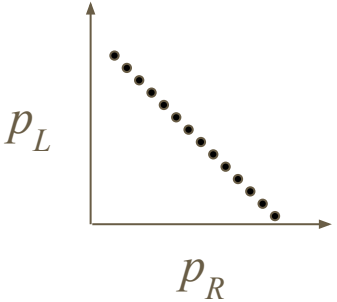




Independent

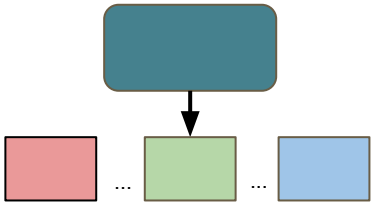


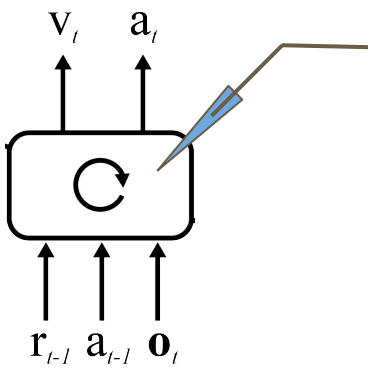
Correlated



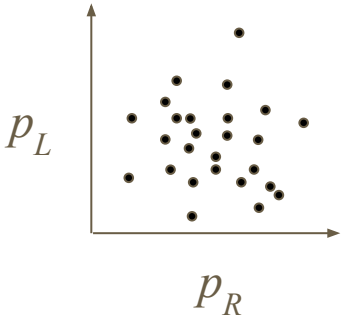
Wang et al. Nature Neuroscience 21 (2018)

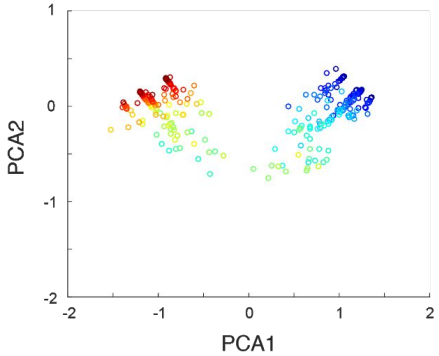
Agent's neural network internalizes task structure



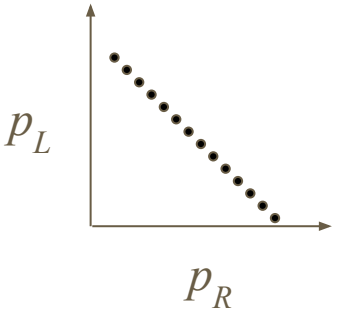


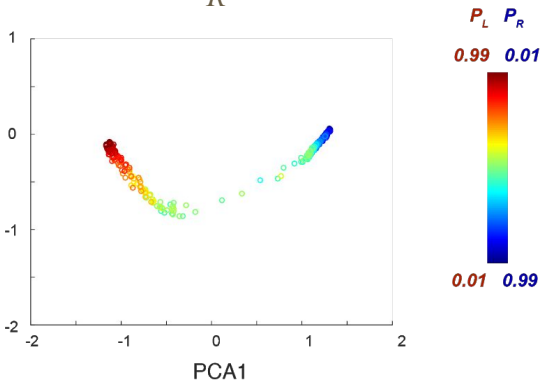
Independent

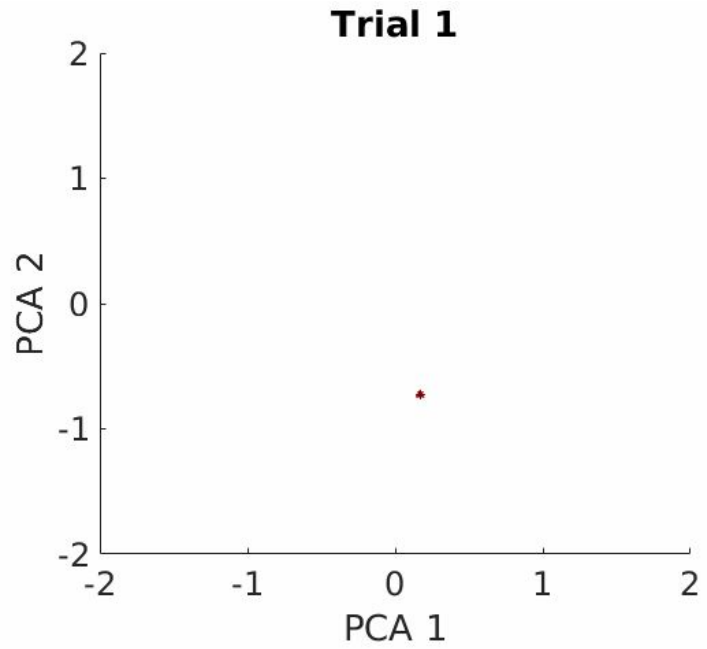


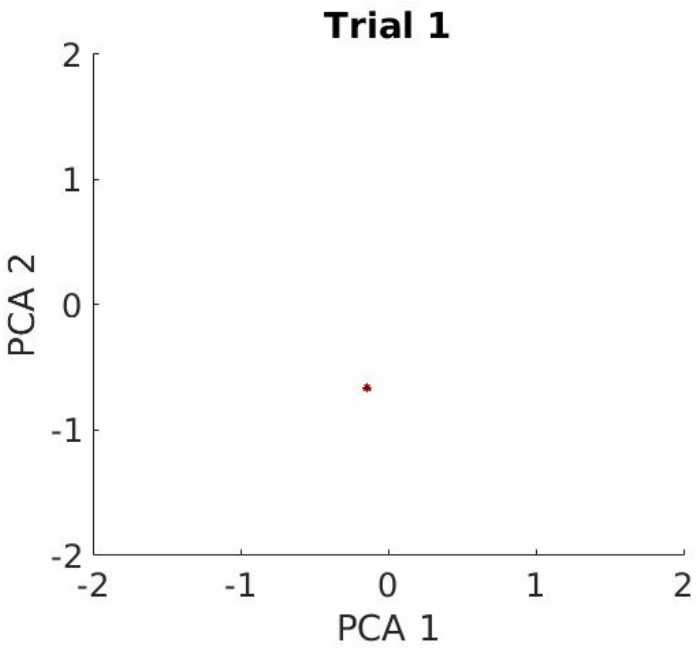


Correlated






Trial 1
PCA 2
PCA 1
2
1
0
-1
-2
-2
-1
0
1
2


Trial 1
PCA 2
PCA 1
2
1
0
-1
-2
-2
-1
0
1
2

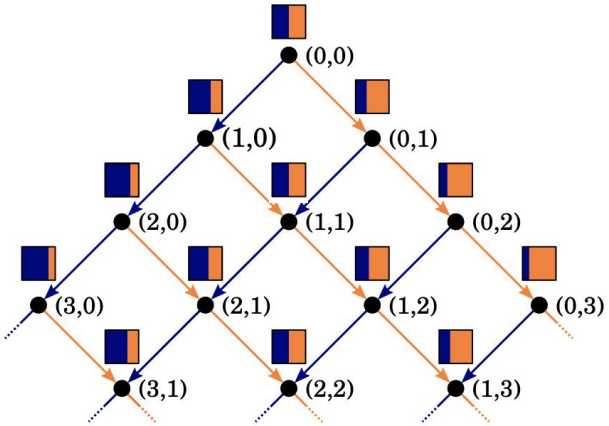
A memory-based meta-learner will necessarily represent task structure

Because of two facts:

- ➔ The meta-learner is trained given observations from a sequence generator with structure, to predict future observations from past history
- ➔ The memory of a meta-learner is limited.

The result is that the meta-learner eventually learns a state representation of sufficient statistics that efficiently captures task structure.

A memory-based meta-learner will necessarily represent task structure



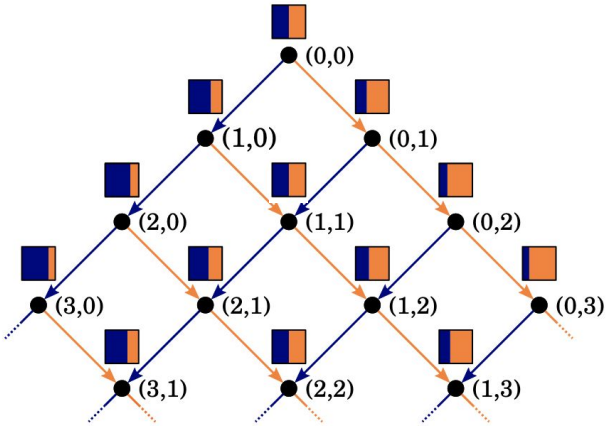
Meta-learning of sequential strategies
Ortega et al, 2019, arXiv:1905.03030

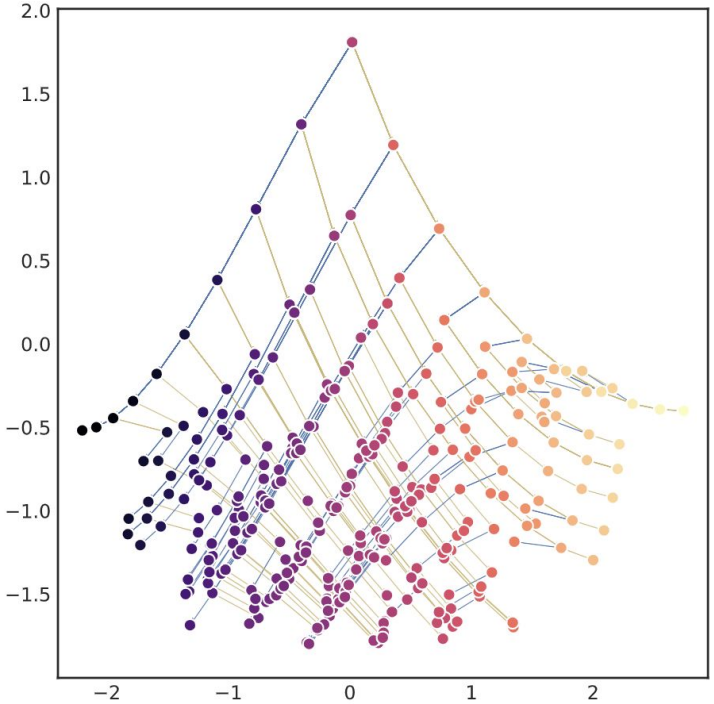
Because of two facts:

- ➔ **The meta-learner is trained given observations from a sequence generator with structure, to predict future observations from past history**
- ➔ **The memory of a meta-learner is limited.**

The result is that the meta-learner eventually learns a state representation of sufficient statistics that efficiently captures task structure.

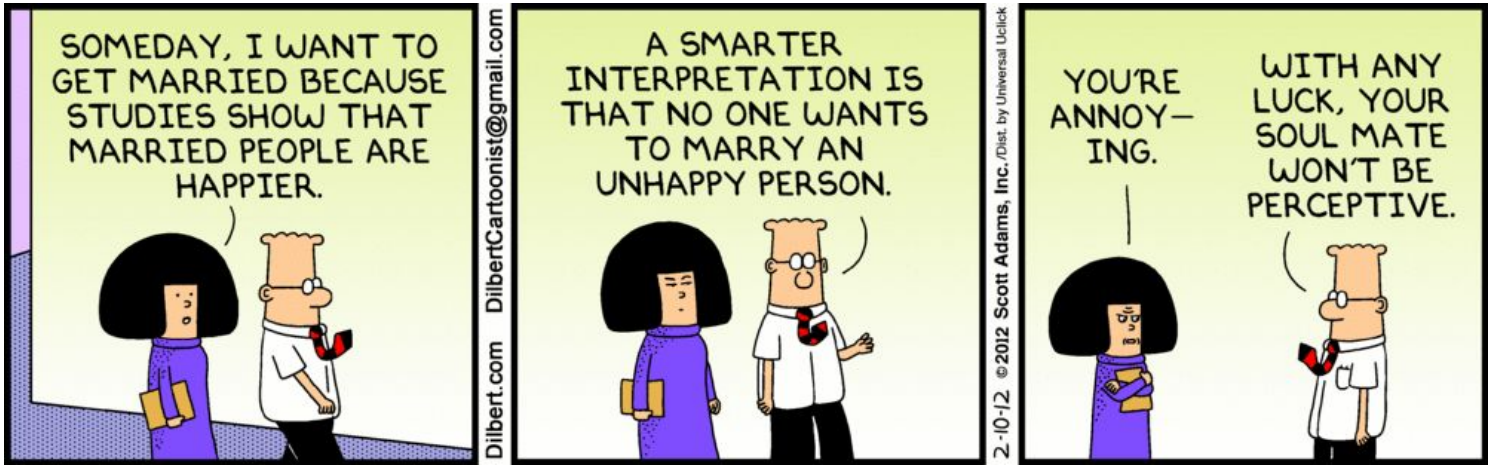
A memory-based meta-learner will necessarily represent task structure

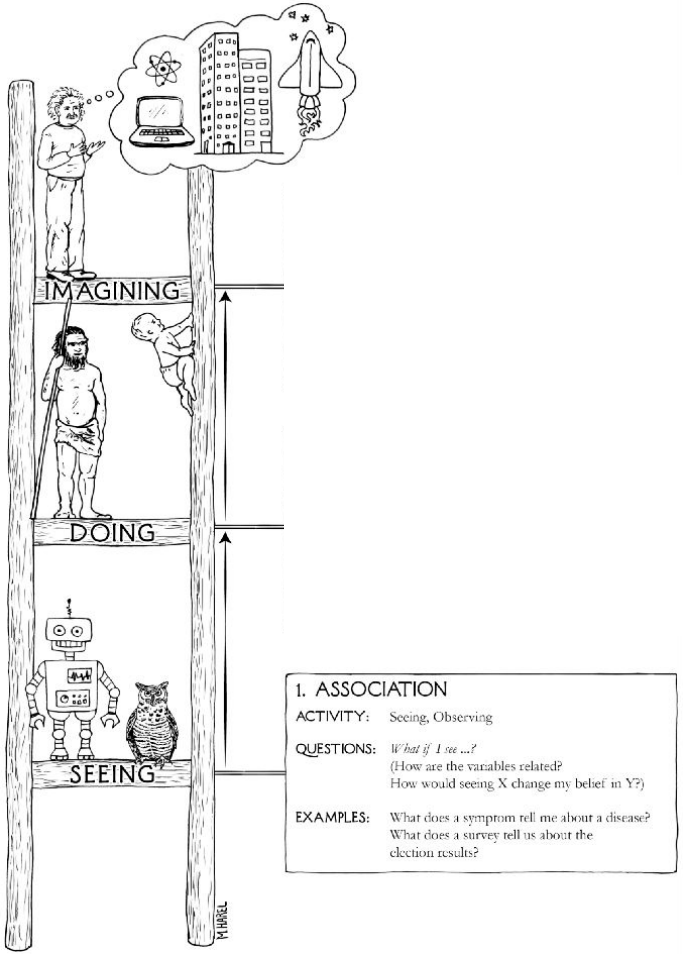


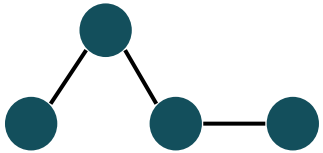


Meta-learning of sequential strategies
Ortega et al, 2019, arXiv:1905.03030

Causally-guided decision-making



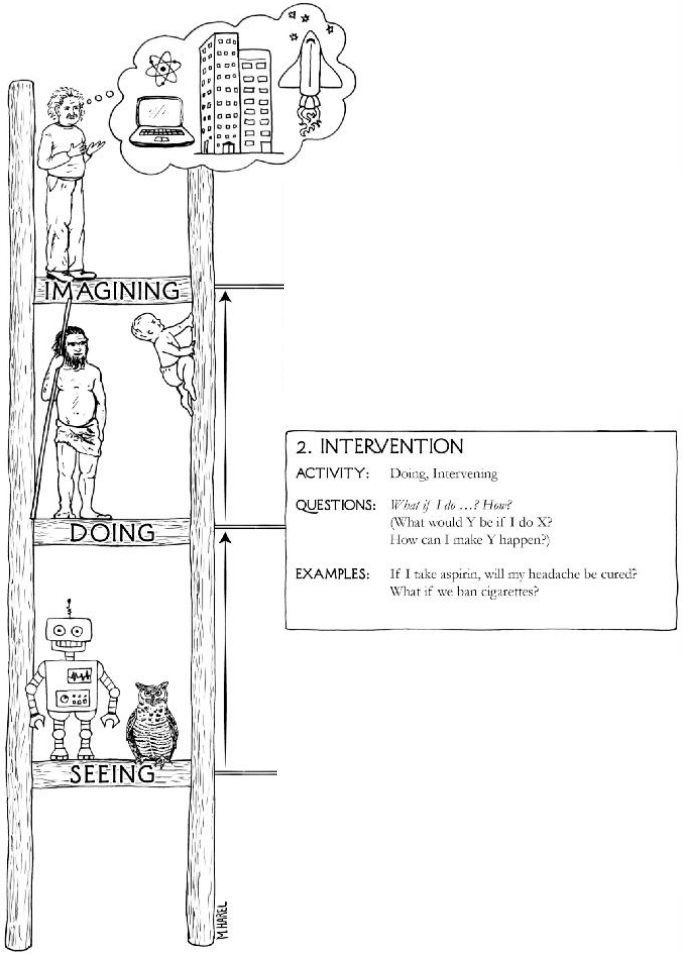


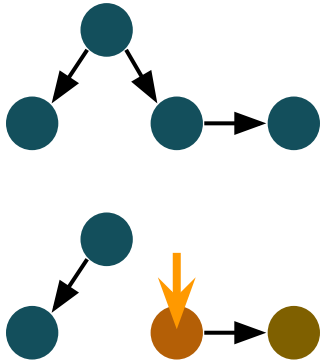


Observing associations, correlations, eg:

“Are drinking wine and having headaches related?”

Judea Pearl's "Ladder of Causation”. Illustrator: Maayan Harel



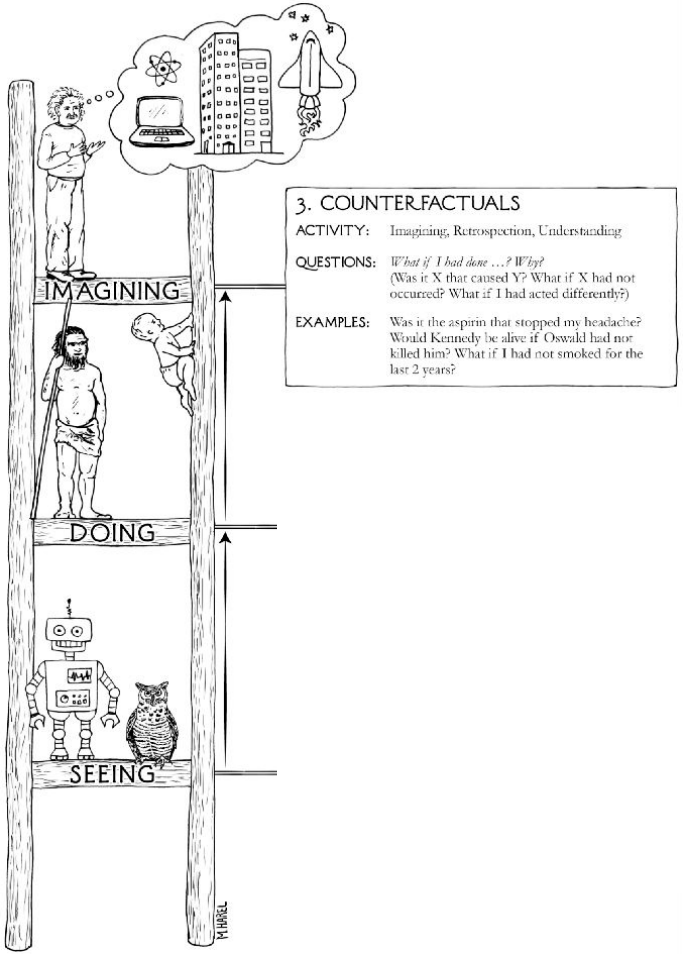


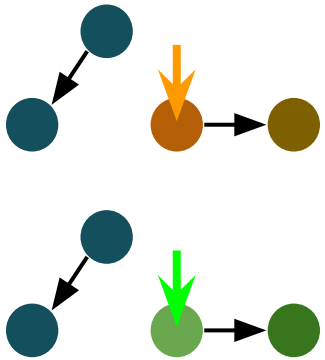
Inferring causal relations from observational data, performing interventions eg:

“If I drink wine, will I get a headache?”

“Does drinking wine cause me to have headaches?”

Judea Pearl's "Ladder of Causation”. Illustrator: Maayan Harel





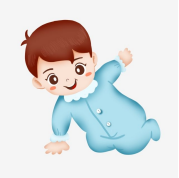
(With same instance-specific noise)

Retrospection, imagining alternatives:

“If I had not drunk wine last night, would I still have a headache?”

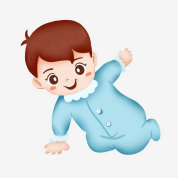
“What if I had drunk soda instead?”

Judea Pearl's "Ladder of Causation”. Illustrator: Maayan Harel



Set developmental trajectory, increasingly optimal causal reasoning from observation

(Piaget et al, Geraci et al, 2011 Dev Science, Schmidt et al, 2011 Plos One, Bonawitz et al, 2010 Cognition, Gopnik et al, 2004 Psych Review)

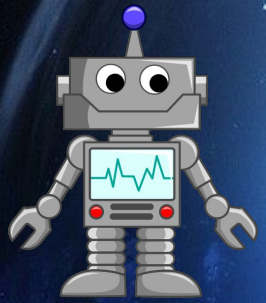


Set developmental trajectory, increasingly optimal causal reasoning from observation

*Ability to perform causal interventions, actively seeking information strategically, individual variability, increased influence of past experience and priors/bias, **apparent** deviation from optimality*

(Gopnik et al, 2001 Dev Psych, Lucas et al, 2013 Cognition, Nussenbaum et al, 2019 psyarxiv, Rehder & Waldman, 2017 Memory & Cognition)

Meta-learning is the DEFAULT, not the exception!



Idea: Meta-learn behaviors that leverage causal knowledge, given structured data and experiences

Idea: Meta-learn behaviors that leverage causal knowledge, given structured data and experiences

Question: Given different types of experience, can agents learn different priors to help it display causal knowledge at different levels?

Idea: Meta-learn behaviors that leverage causal knowledge, given structured data and experiences

Question: Given different types of experience, can agents learn different priors to help it display causal knowledge at different levels?

Approach:

- Set up tasks that allow our agents to demonstrate causal strategies, under different task requirements
- Implement various controls, comparing against non-learning benchmarks, testing on held-out graphs and interventions
- Detailed interrogation of behavior

Idea: Meta-learn behaviors that leverage causal knowledge, given structured data and experiences

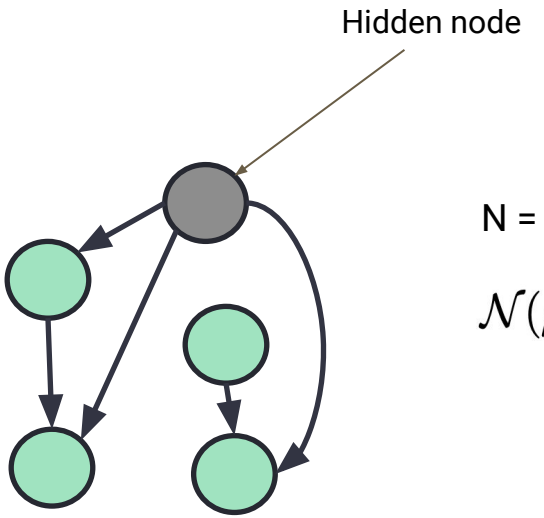
Question: Given different types of experience, can agents learn different priors to help it display causal knowledge at different levels?

Approach:

Type of experience	Type of inference
~~Observational~~	~~Causal inference~~
Interventional	Confounder resolution
Noise information	Counterfactual

An example episode

Environment



N = 5

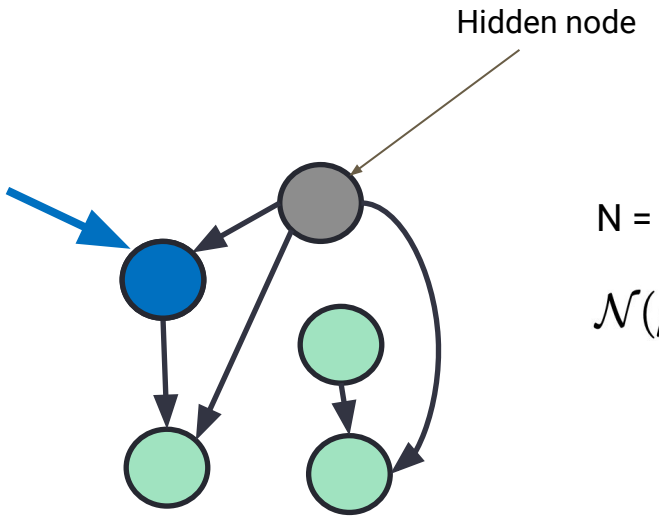
$$\mathcal{N}(\mu = \sum_j w_{ji} X_j, \sigma = 0.1)$$

Dasgupta, et al. 2019, arXiv:1901.08162

An example episode

Environment

Interactions *(N-1 steps)*



N = 5

$$\mathcal{N}(\mu = \sum_j w_{ji} X_j, \sigma = 0.1)$$

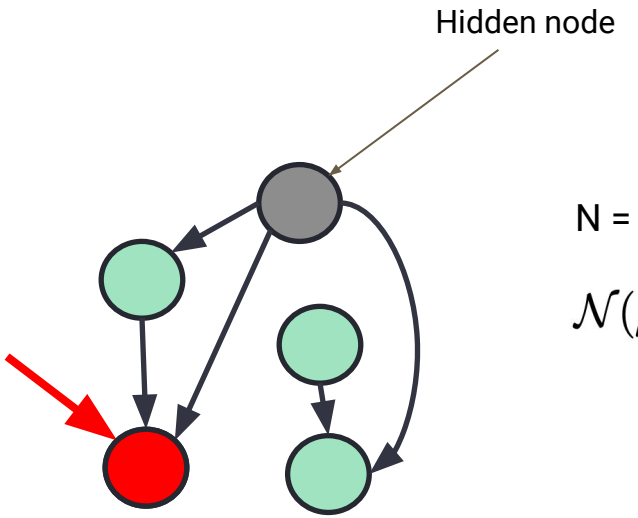
Dasgupta, et al. 2019, arXiv:1901.08162

An example episode

Environment

Interactions *(N-1 steps)*

Previously unobserved event on test step



N = 5

$$\mathcal{N}(\mu = \sum_j w_{ji} X_j, \sigma = 0.1)$$

Dasgupta, et al. 2019, arXiv:1901.08162

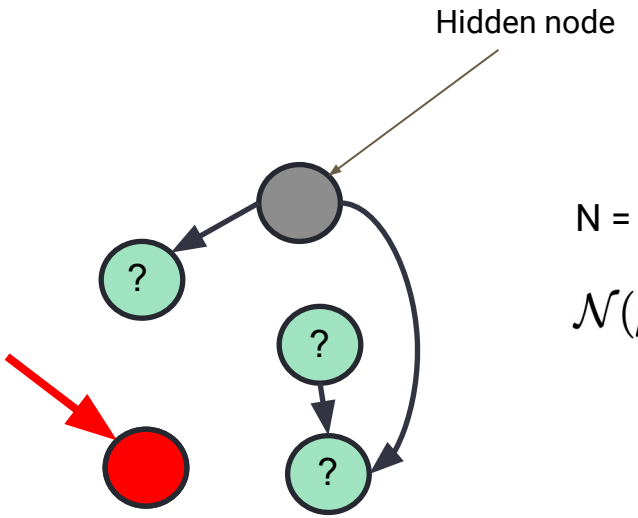
An example episode

Environment

Interactions *(N-1 steps)*

Previously unobserved event on test step

Predict highest value node

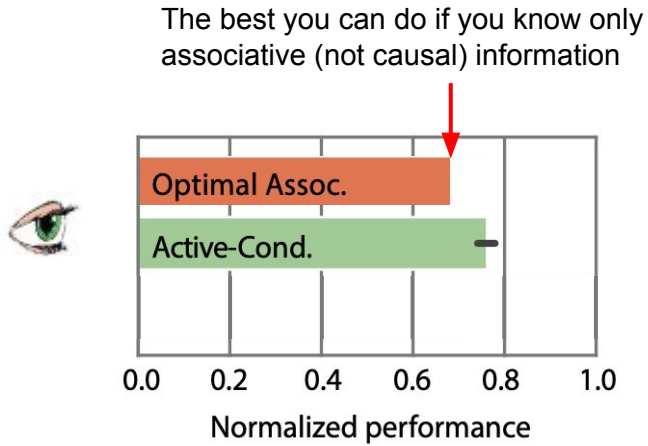


N = 5

$$\mathcal{N}(\mu = \sum_j w_{ji} X_j, \sigma = 0.1)$$

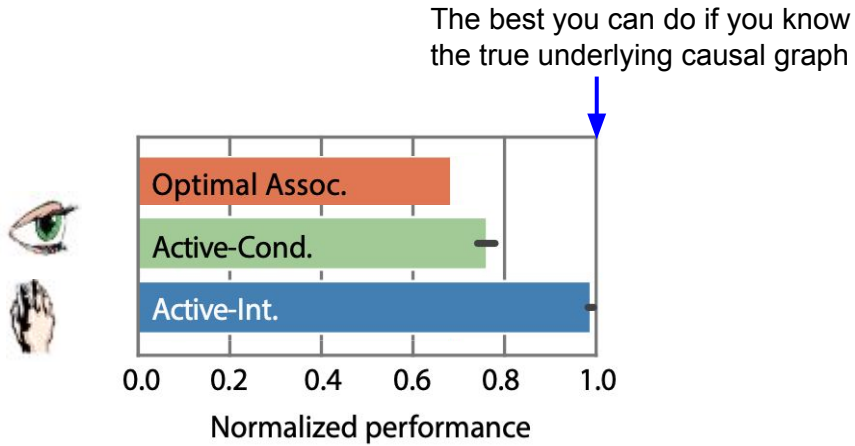
Dasgupta, et al. 2019, arXiv:1901.08162

Meta-RL agent learns to perform interventions at a performance close to ceiling (best you can do given knowledge of ground truth causal graph).



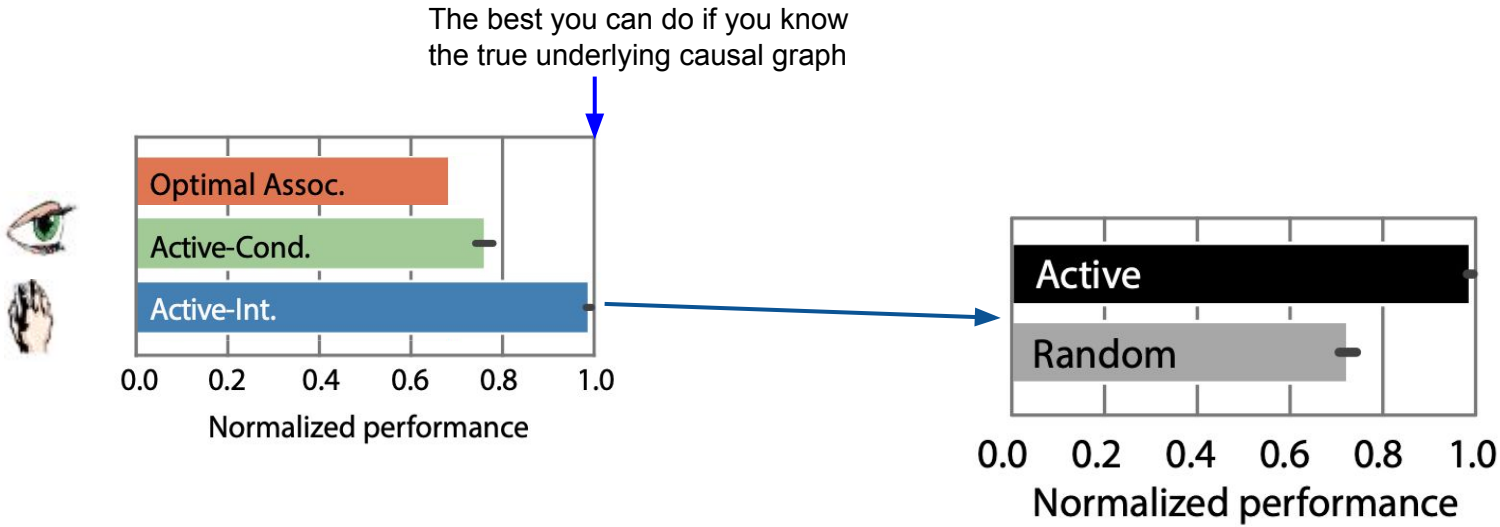
Dasgupta, et al. 2019, arXiv:1901.08162

Meta-RL agent learns to perform interventions at a performance close to ceiling (best you can do given knowledge of ground truth causal graph).



Dasgupta, et al. 2019, arXiv:1901.08162

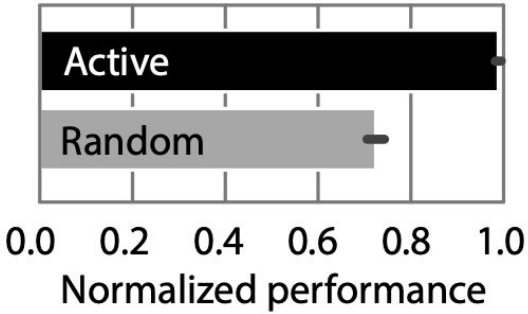
Performs better than an agent that cannot choose which node it gets to intervene on.

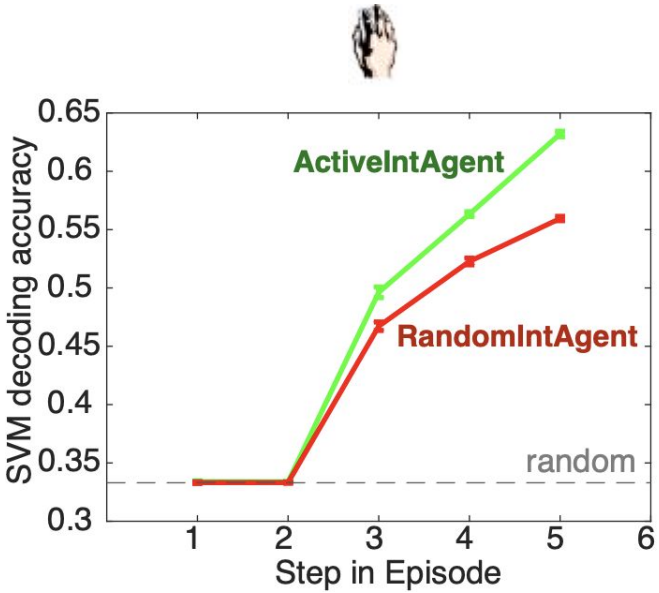


Dasgupta, et al. 2019, arXiv:1901.08162

Active interventional policy allows agents to more accurately encode ground truth causal graph in hidden state, vs random interventions.

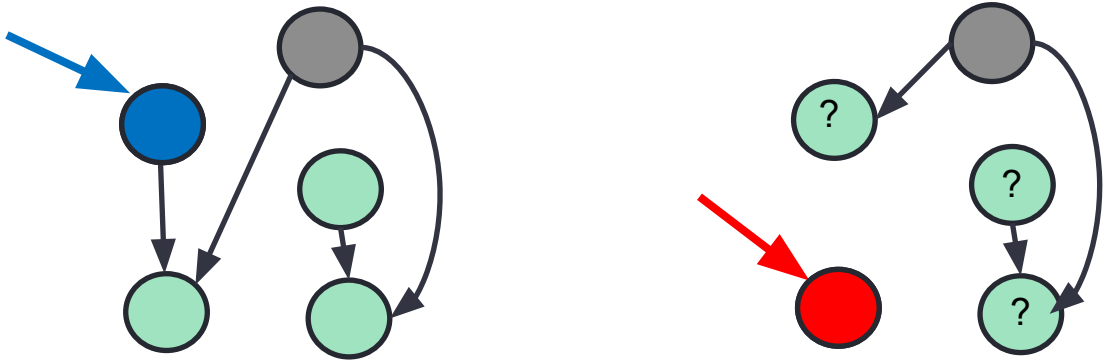
Note: doesn't need to fully represent the graph in order to perform at ceiling - since it isn't necessary for the task.





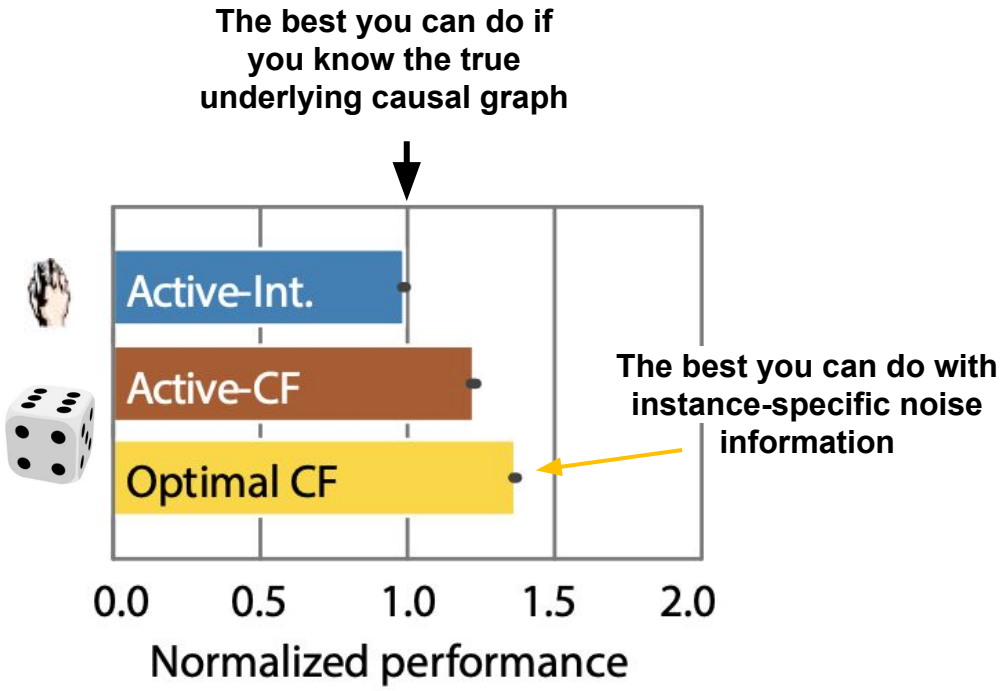
Dasgupta, et al. 2019, arXiv:1901.08162

Learning from instance-specific info (counterfactuals)

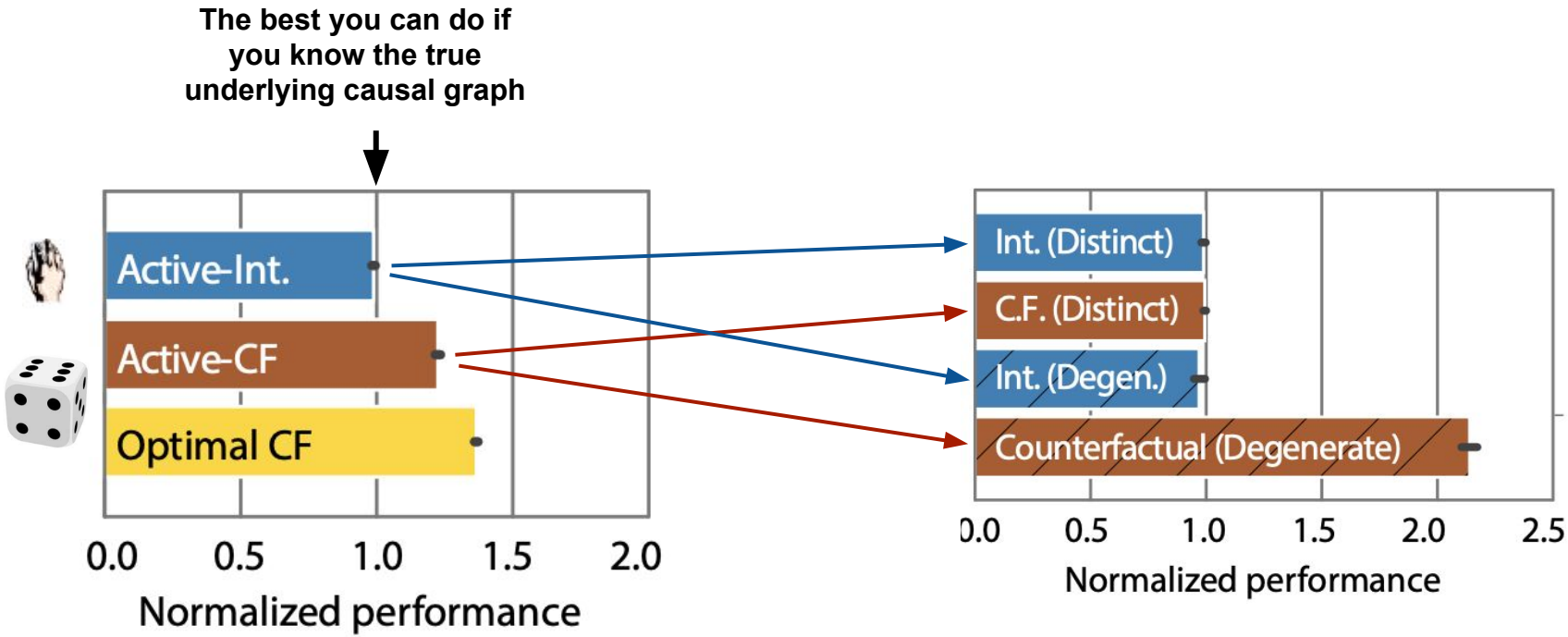


$$\mathcal{N}(\mu = \sum_i w_{ji} X_j, \sigma = 0.1)$$

Dasgupta, et al. 2019, arXiv:1901.08162



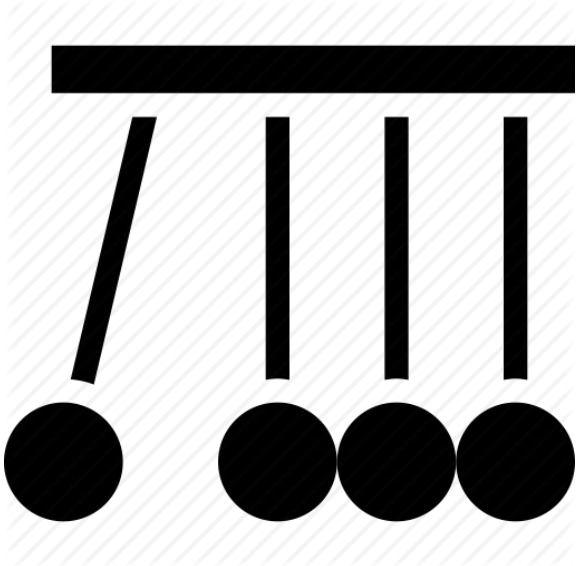
Dasgupta, et al. 2019, arXiv:1901.08162



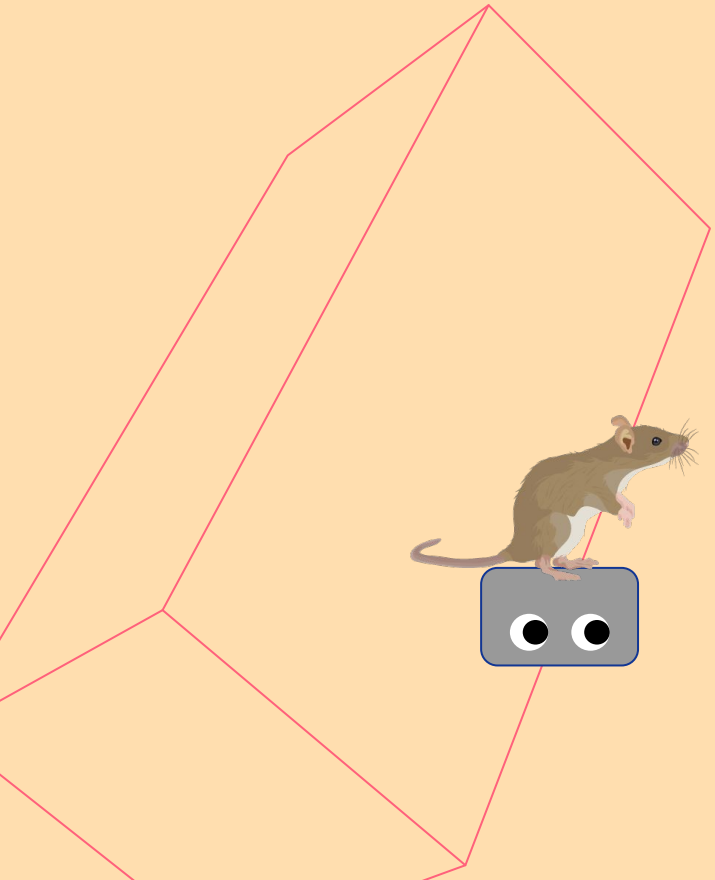
Dasgupta, et al. 2019, arXiv:1901.08162

Implications

- Within a meta-RL setup, agents are apparently capable of acting to acquire and use causal information for better task performance.
- Assumes that we have the right representations, next challenge will be to combine with deep learning modules to learn these representations
- Performance and behavior is experience-dependent (meta-learned from the data), so task-design is crucial.



DeepMind



With many thanks to:

Ishita Dasgupta (Harvard)
Matt Botvinick (DeepMind)
Zeb Kurth-Nelson (DeepMind)
Kevin Miller (DeepMind)
Pedro Ortega (DeepMind)
Silvia Chiappa (DeepMind)
...and countless colleagues at DeepMind

Questions?