DeepMind

Royal Society meeting:
Beyond the symbols vs signals debate

Meta-learning as bridging the neuro-symbolic gap in AI

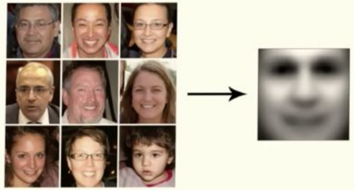
Jane Wang
Google DeepMind
28–29 October 2024

Two types of abstraction / generalization

The two poles of abstraction: type 1 vs type 2

Prototype-centric (value-centric) abstraction

- Set of prototypes + distance function
  - Example: classify face vs. non-face using abstract features
- Abstract wrt details not present in the prototypes
- Obtained by clustering concrete samples into prototypes
  - This is a value analogy!



Program-centric abstraction

- Graph of (usually discrete) operators where input nodes can take different values within a type
  - Example: function that sorts a list
- Abstract wrt input nodes values
- Obtained by merging specialized functions under a new abstract signature
  - This is a program analogy!

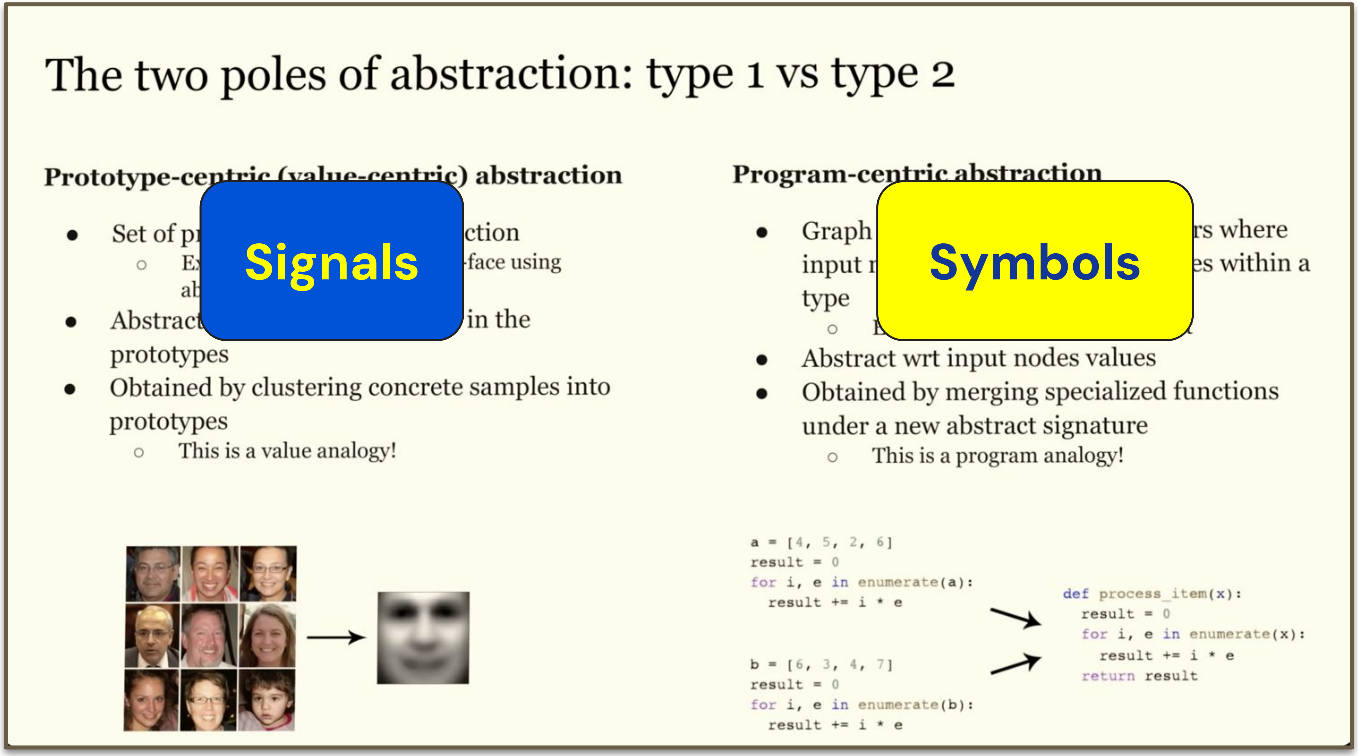
```
a = [4, 5, 2, 6]
result = 0
for i, e in enumerate(a):
  result += i * e

b = [6, 3, 4, 7]
result = 0
for i, e in enumerate(b):
  result += i * e
```

→

```
def process_item(x):
  result = 0
  for i, e in enumerate(x):
    result += i * e
  return result
```

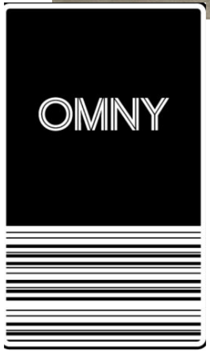
Francois Chollet, keynote talk at AGI-24

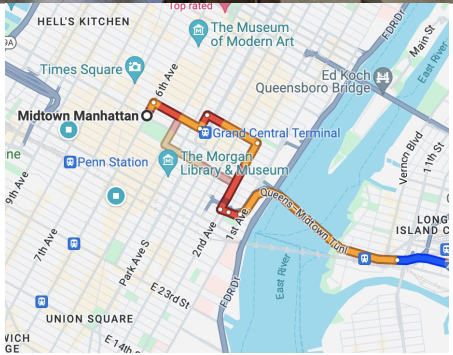
Two types of abstraction / generalization

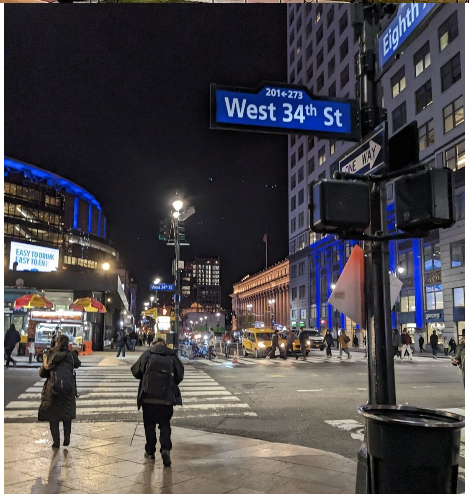


Francois Chollet, keynote talk at AGI-24

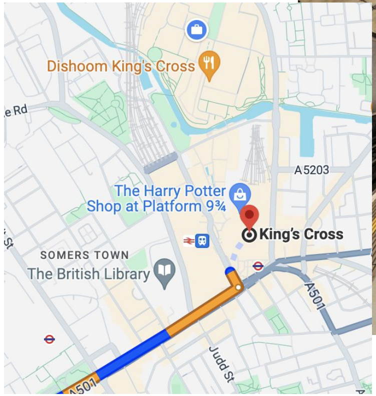
The ability to navigate and act within new environments...







Depends on prior, structured experience



Solving tasks requires mapping to the right representations

Inside subway station

Central square

Payment card

Location on map

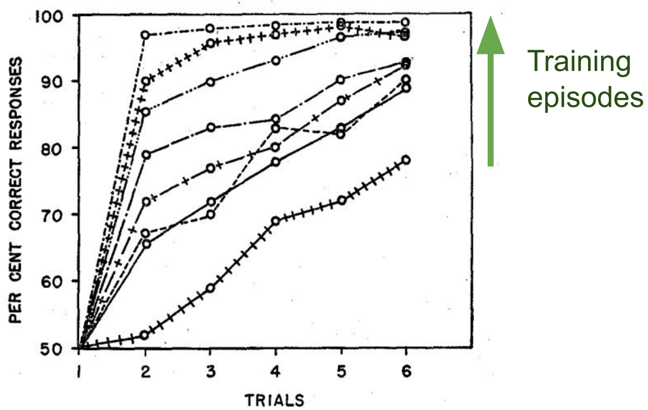
Platform

Front of station

“Meta-learning” is one method of learning useful representations

A process of learning priors or useful representations from previous experience to enable faster learning or better decisions

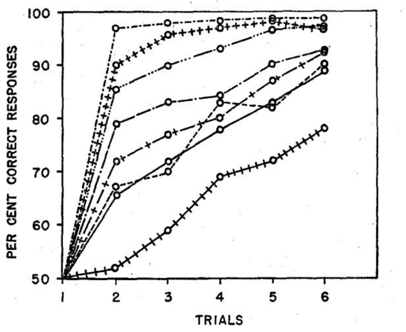
The “Harlow task”

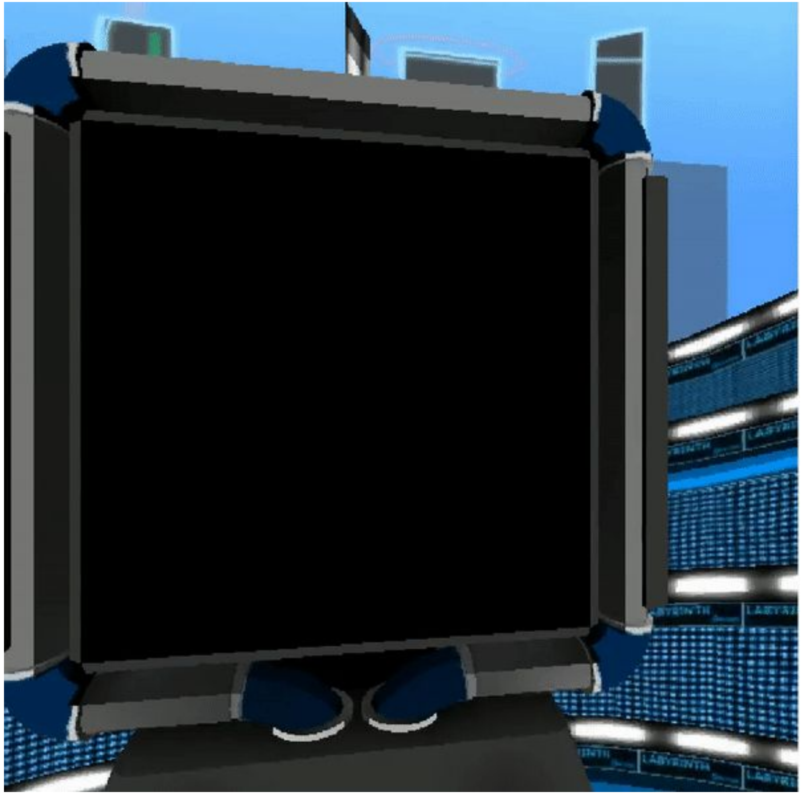


Harlow, 1949, Psychological Review

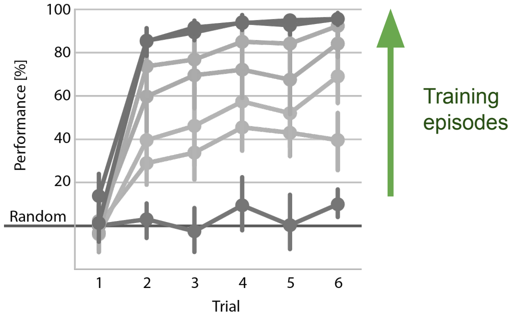
Artificial agent (weights frozen)

Animal



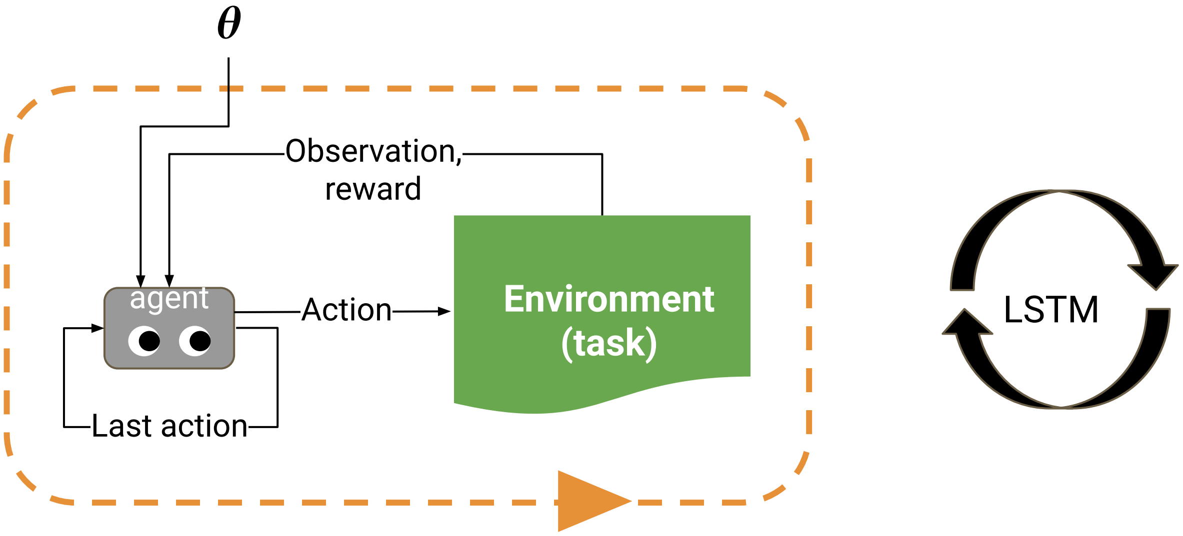


Artificial agent

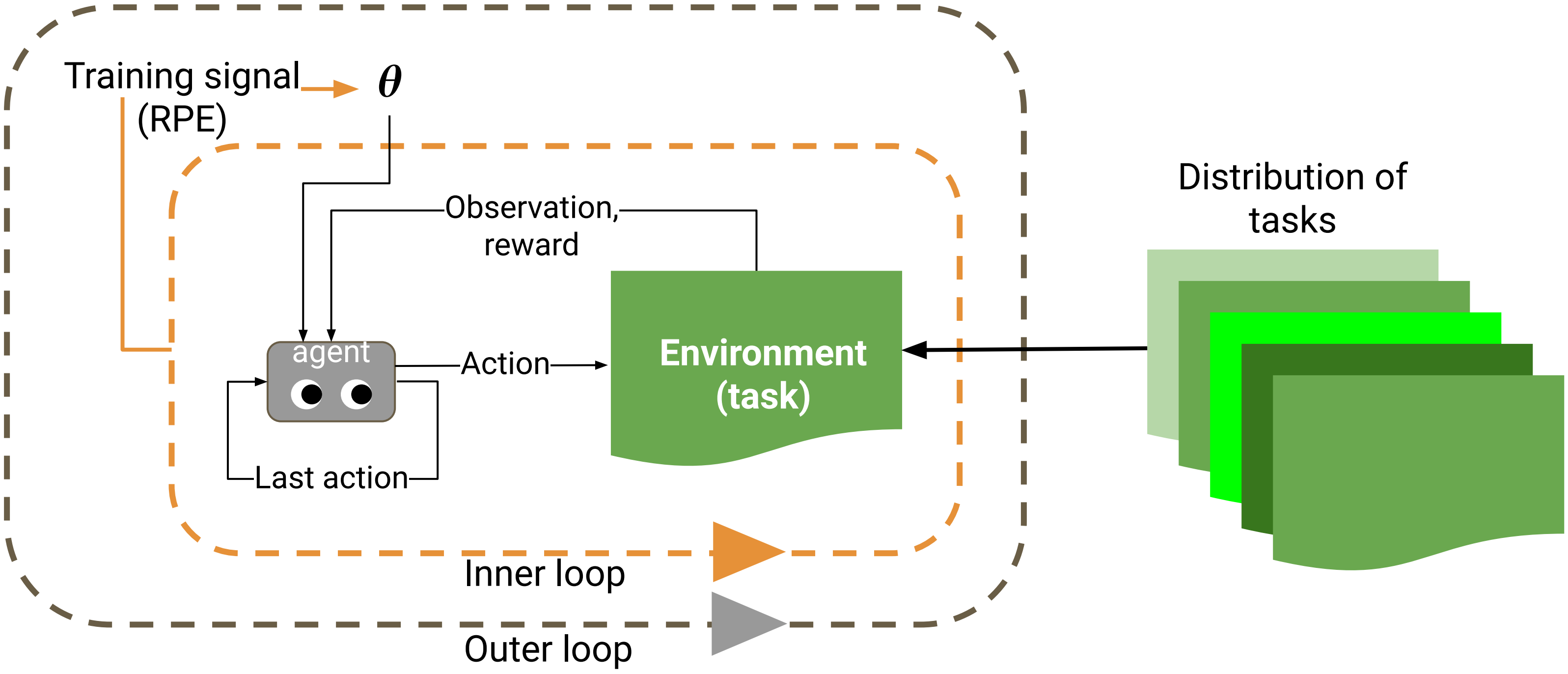


Wang, Kurth-Nelson, et al. Nature Neuroscience (2018)

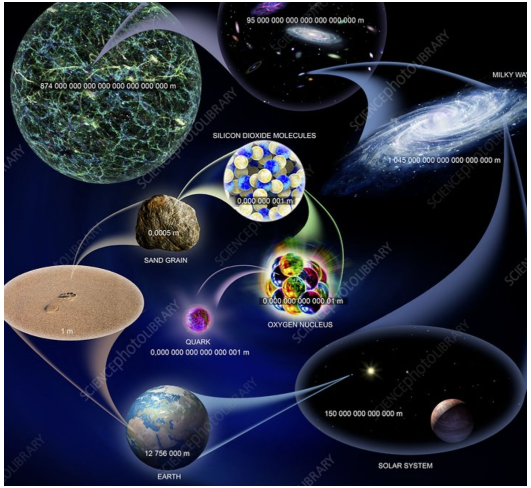
Meta-reinforcement learning



Meta-reinforcement learning



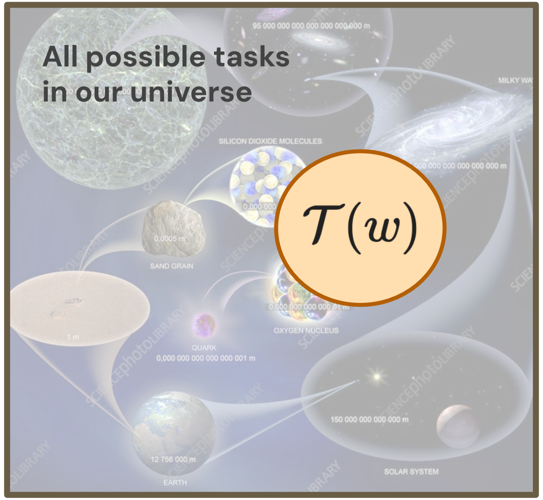
What is structured training data?



In meta-learning, we only consider a subset of tasks of interest

Let's assume that all training tasks are sampled from a generative process with latent parameters $\boldsymbol{w}$ that generates a sequence of observations, conditioned on past states and actions

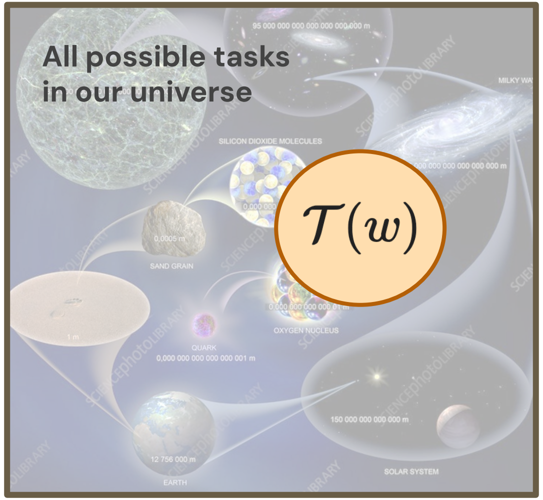
e.g. rules of physics, initial conditions, causal relationships, variables of interest, etc (note that N($\boldsymbol{w}$) << N(tasks))



Tasks determine the states we can reach, with different utilities

Let's assume that all training tasks are sampled from a generative process with latent parameters $\boldsymbol{w}$ that generates a sequence of observations, conditioned on past states and actions

e.g. rules of physics, initial conditions, causal relationships, variables of interest, etc (note that N($\boldsymbol{w}$) << N(tasks))

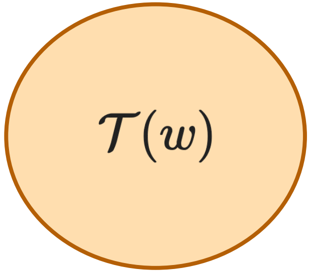


Every episode *i*, sample a task $T_i \sim \mathcal{T}(w)$

This task determines how an agent can interact with the environment to get to states with different utility or reward

Generative process with latent parameters ***w*** that generates task

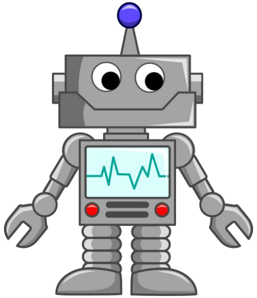
$$T_i \sim \mathcal{T}(w)$$



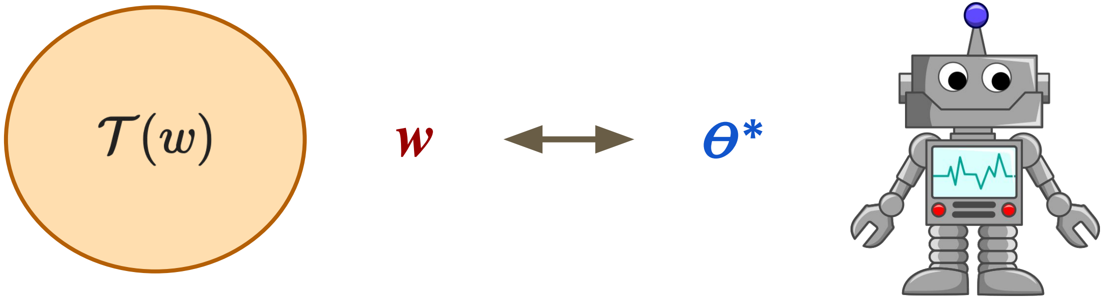
We train a policy (agent) with learned parameters ***θ*** that interacts with sampled task T_i to maximize expected total utility U (or minimize loss) for all the states s visited in task T

$$U(T_i; \pi_\theta) = \sum_t u(s_{t,i}; \pi_\theta)$$

$$\max_\theta U(T_i; \pi_\theta)$$



Every episode *i* (gradient update), sample a task. Repeat for many $T_i \sim \mathcal{T}(w)$

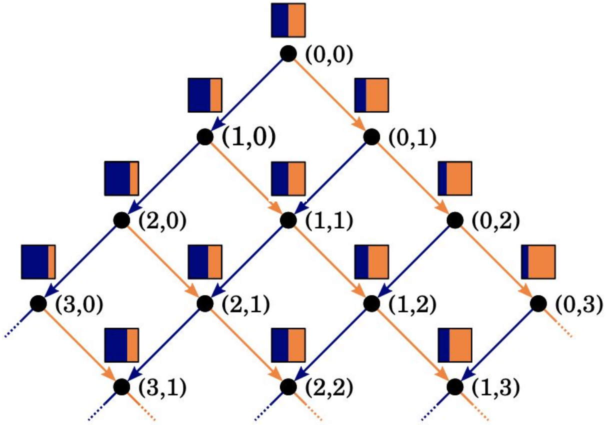


If trained to optimal θ^*, meta-learning parameters will represent sufficient statistics for this generative process

For more formal description: see Ortega et al, 2019. Meta-learning of sequential strategies, arXiv:1905.03030

A memory-based meta-learner will converge to represent task structure

w →

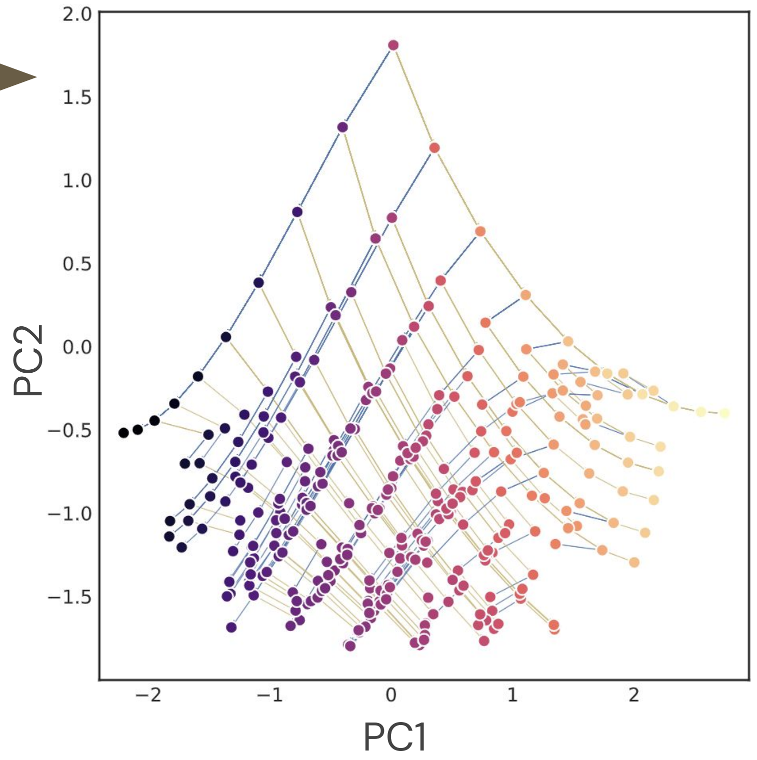


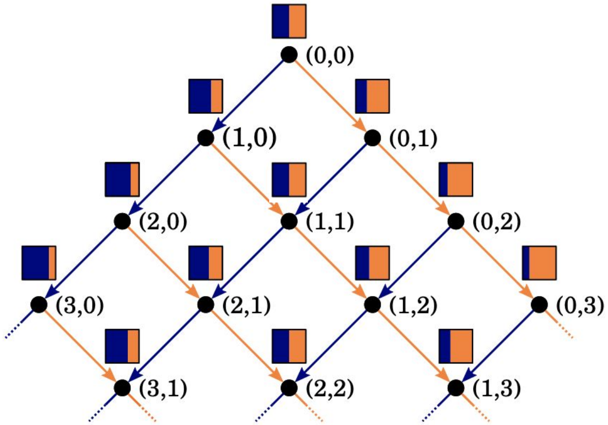
Meta-learning of sequential strategies
Ortega, Wang, et al, 2019, arXiv:1905.03030

A memory-based meta-learner will converge to represent task structure

w →

$f(x,\theta)$ →

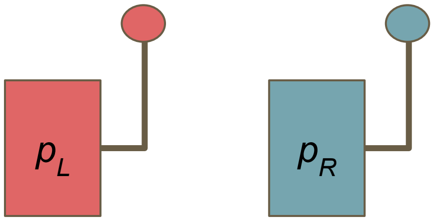




Meta-learning of sequential strategies
Ortega, Wang, et al, 2019, arXiv:1905.03030

2-armed bandits
independently drawn from
uniform Bernoulli distribution

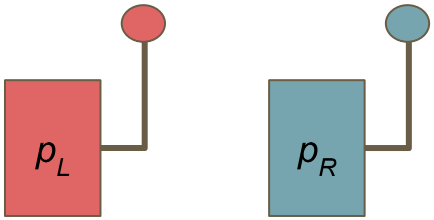
Held constant for 100 trials
=1 episode



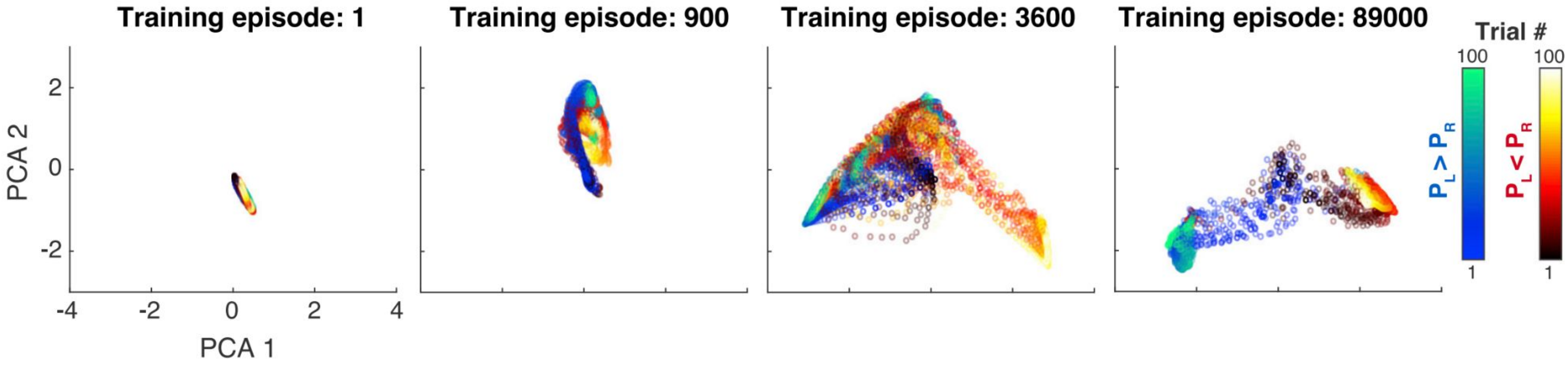
p_i = probability of payout,
drawn uniformly from [0,1],

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],



Botvinick, Wang, et al, 2020. Deep reinforcement learning and its neuroscientific implications. Neuron

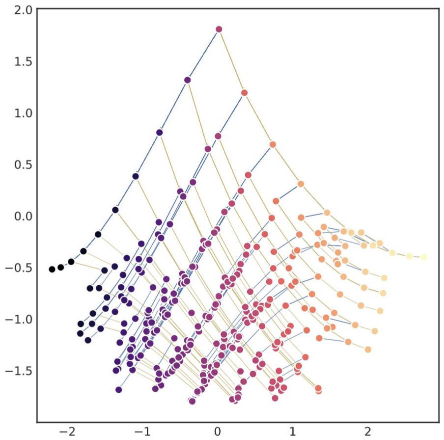
A good (useful) representation provides a mapping between raw sensory data and the underlying task-relevant variables of the set of tasks

The end result of meta-learning is the acquisition of this representation

w →

OR

$f(x,\theta)$

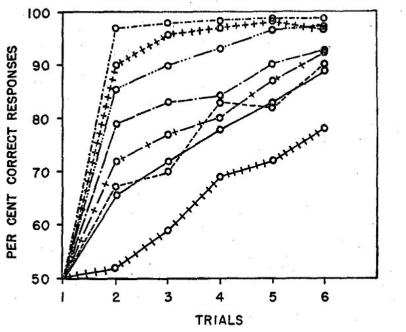


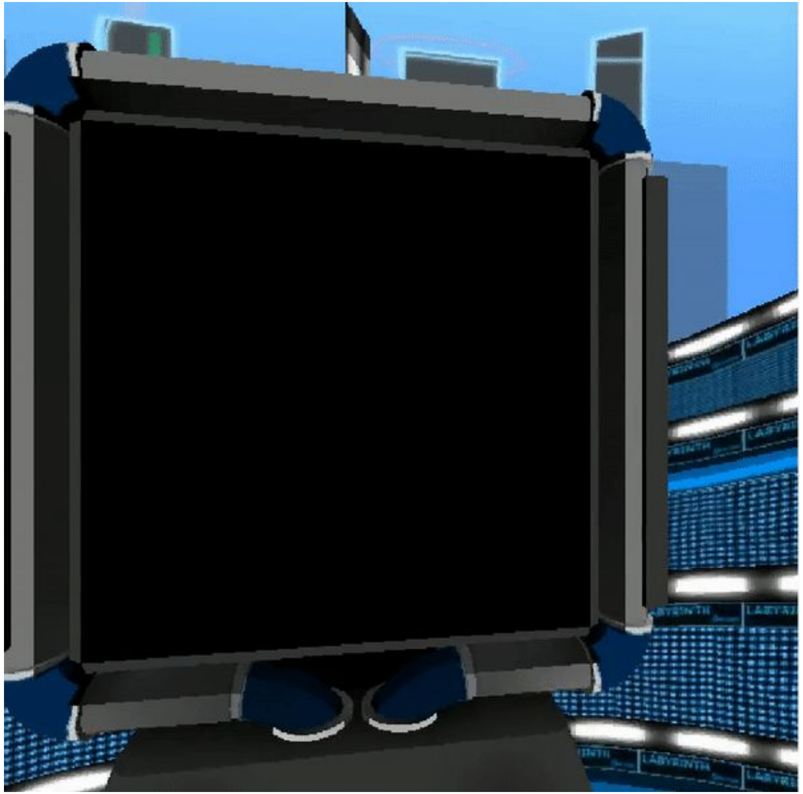
A good (useful) representation provides a mapping between raw sensory data and the underlying task-relevant variables of the set of tasks

The end result of meta-learning is the acquisition of this representation

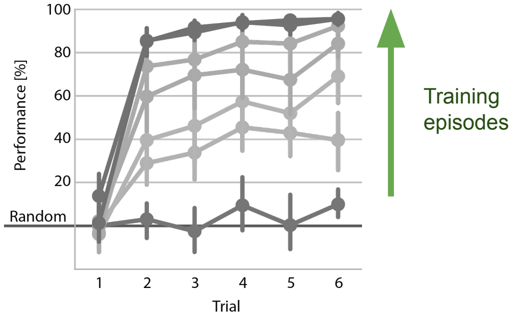
BUT there is no guarantee that this representation can be learned in any reasonable amount of time!

Animal

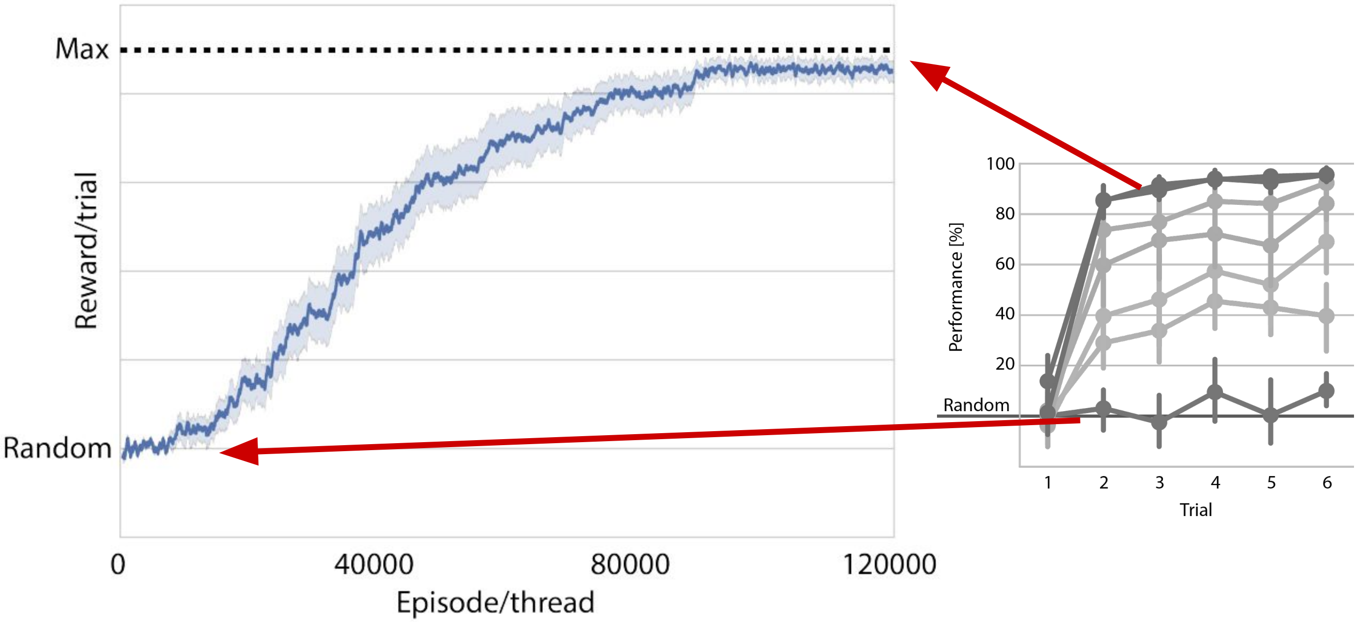




Artificial agent

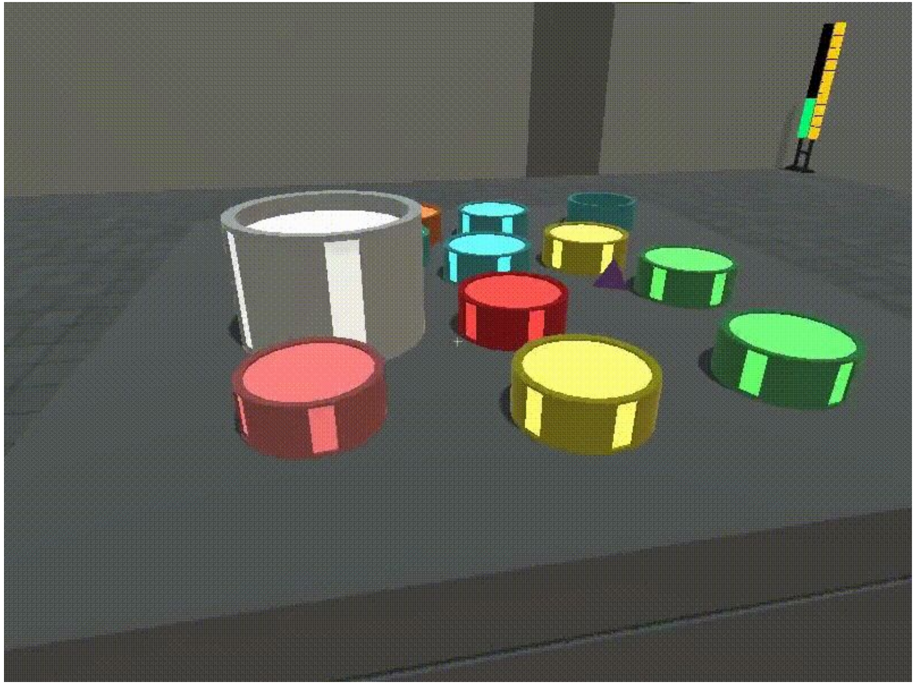


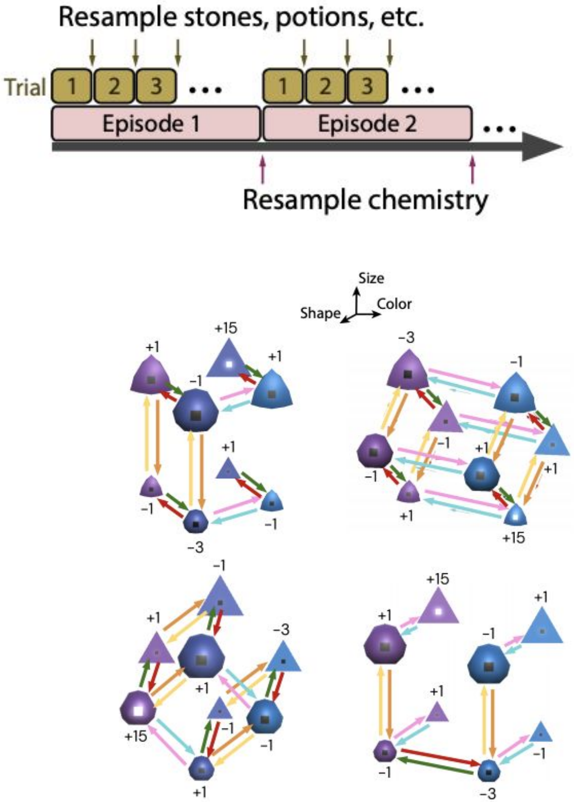
Wang, Kurth-Nelson, et al. Nature Neuroscience (2018)


Max
Random
Reward/trial
0
40000
80000
120000
Episode/thread
100
80
60
40
20
Random
Performance [%]
1
2
3
4
5
6
Trial

(32 threads)

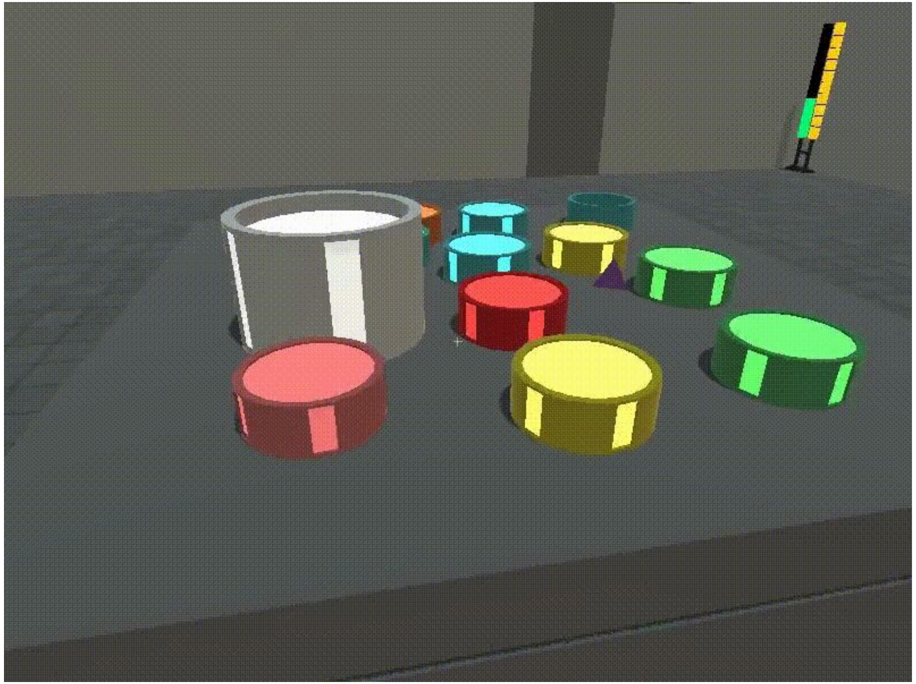
Alchemy: A meta-reinforcement learning benchmark



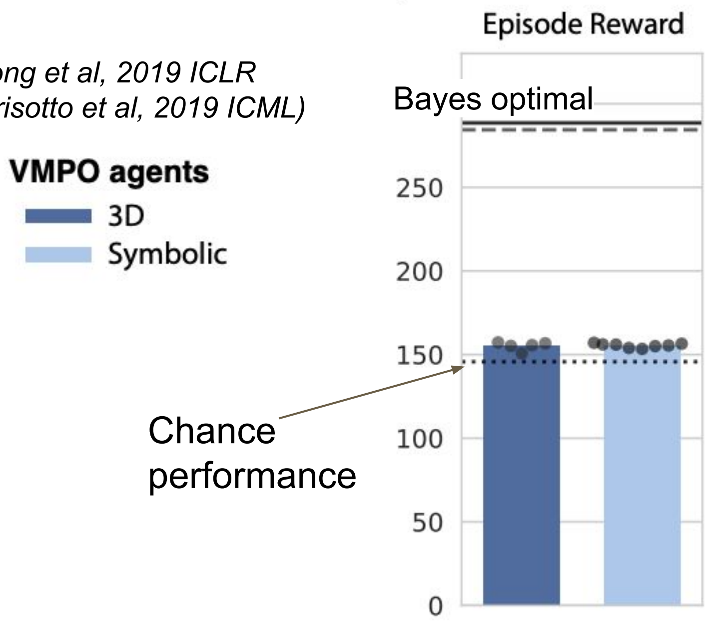


Wang, King, et al, 2021, NeurIPS Datasets and Benchmarks

Alchemy: A meta-reinforcement learning benchmark

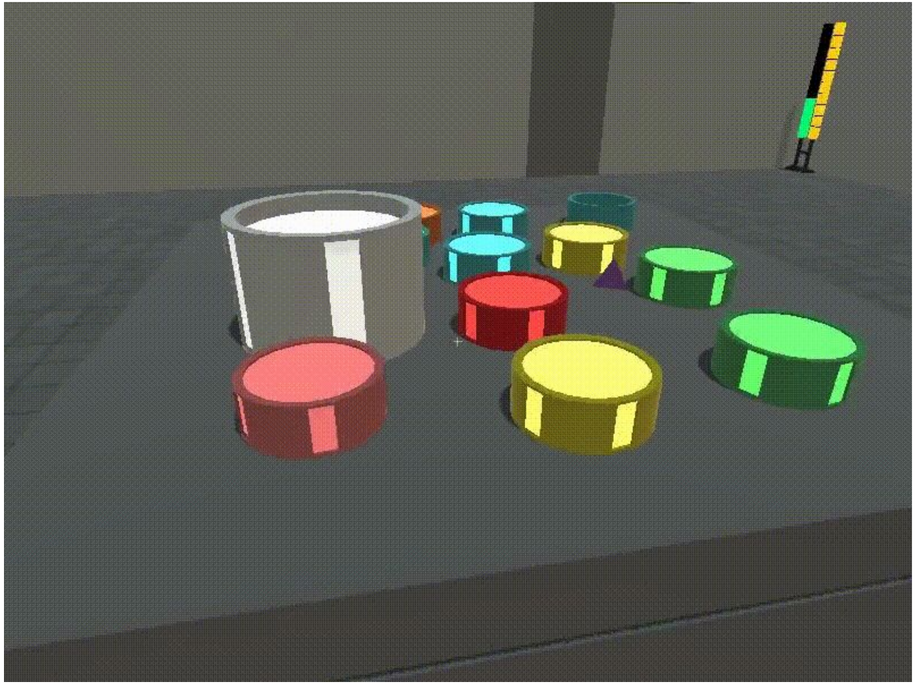


(Song et al, 2019 ICLR
Parisotto et al, 2019 ICML)

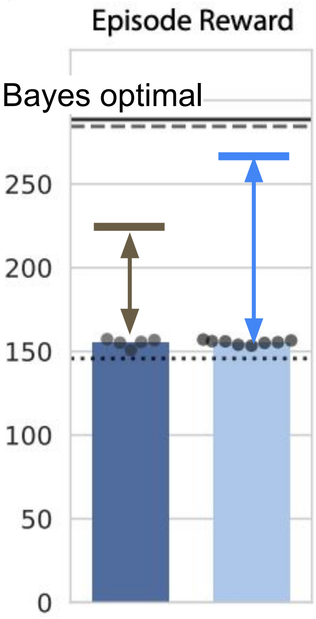


Wang, King, et al, 2021, NeurIPS Datasets and Benchmarks

Alchemy: A meta-reinforcement learning benchmark



Improvement when training with auxiliary task specifically designed to give the right task-related representations



Wang, King, et al, 2021, NeurIPS Datasets and Benchmarks

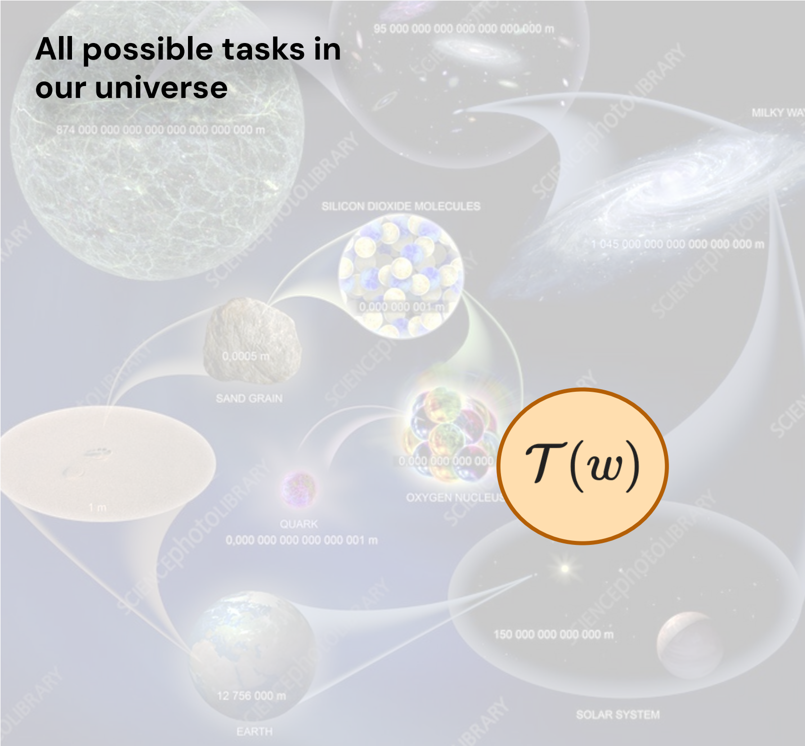
On the role of planning in model-based deep reinforcement learning

Jessica B. Hamrick,* Abram L. Friesen, Feryal Behbahani, Arthur Guez, Fabio Viola, Sims Witherspoon, Thomas Anthony, Lars Buesing, Petar Veličković, Théophane Weber*
DeepMind, London, UK

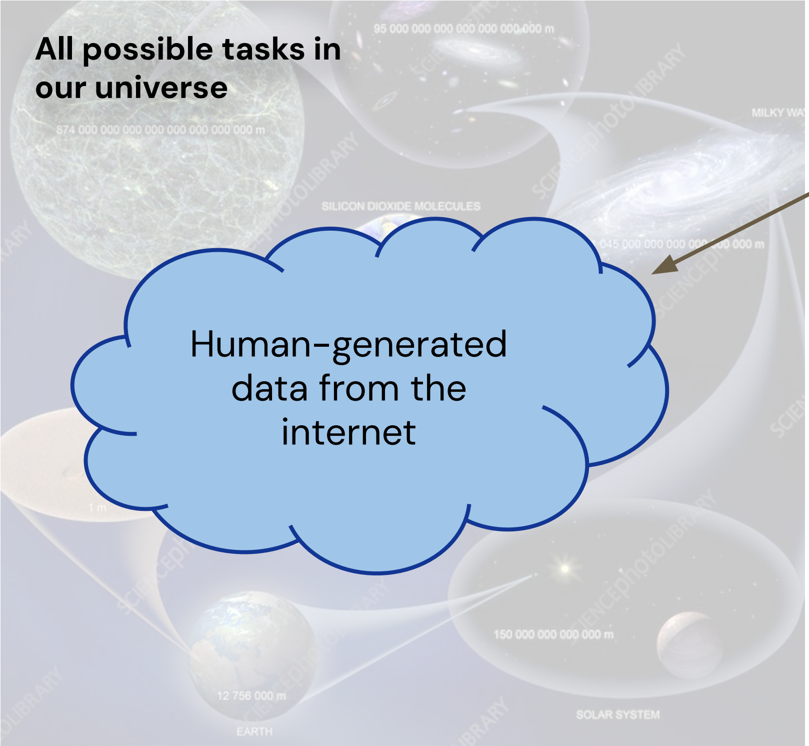
- ➔ Looked at the role that planning plays in **generalization** for agents like MuZero
- ➔ Found it was more important to have the **right abstractions and representations** in the value and policy than learning a correct model of the environment or doing extensive planning
- ➔ How do we get our models to learn these right representations?

Hamrick et al, 2021, ICLR

Answer(?): first pre-train on real-world data at scale



Answer(?): first pre-train on real-world data at scale

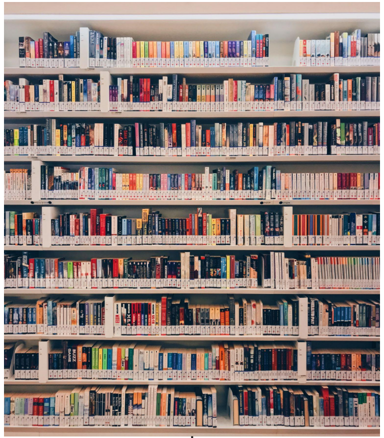


First train on this

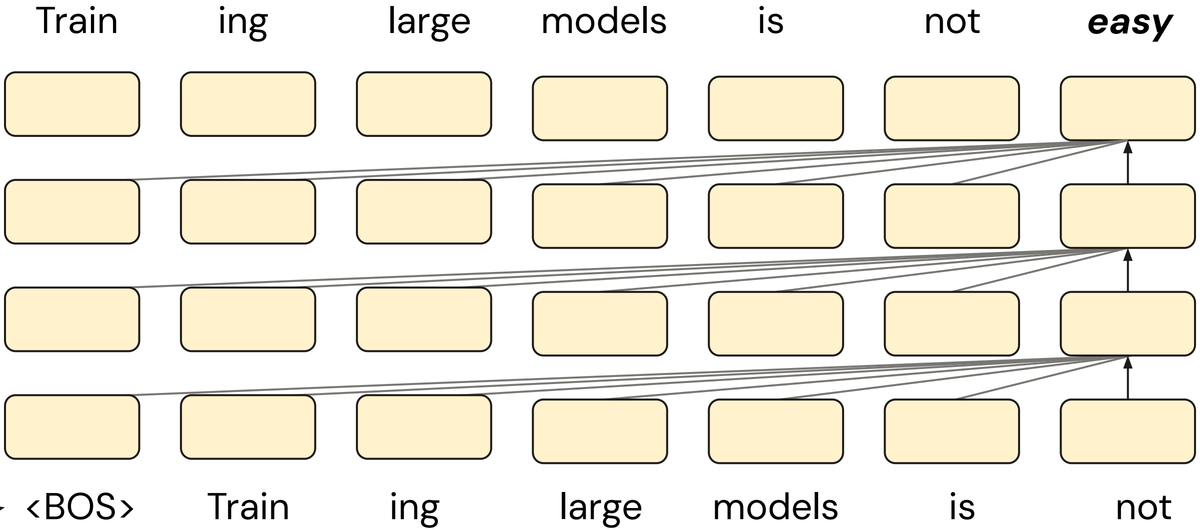
And then fine tune, few-shot, or even zero-shot prompt on this

$\mathcal{T}(w)$

Language model training is not active; they passively predict the next token in someone else's language



A lot of text on the internet, written by some humans



In practice, it does surprisingly well (not human-level)

- LMs provide useful priors for causal reasoning mechanisms, e.g. for identifying causal structures from data

- LMs can be prompted to interactively use tools (e.g. APIs) to achieve a task

Causal Reasoning and Large Language Models: Opening a New Frontier for Causality

Emre Kıcıman*
Microsoft Research
emrek@microsoft.com

Robert Ness
Microsoft Research
robertness@microsoft.com

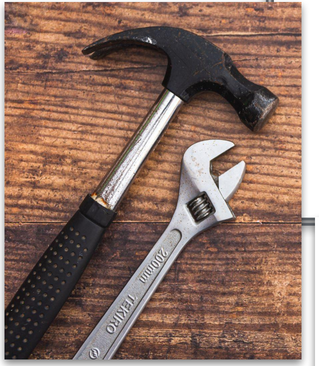
Amit Sharma
Microsoft Research
amshar@microsoft.com

Chenhao Tan
University of Chicago
chenhao@uchicago.edu

Toolformer: Language Models Can Teach Themselves to Use Tools

Timo Schick Jane Dwivedi-Yu Roberto Dessì[†] Roberta Raileanu Maria Lomeli Luke Zettlemoyer Nicola Cancedda Thomas Scialom

Meta AI Research [†]Universitat Pompeu Fabra



Chat Plugins Beta

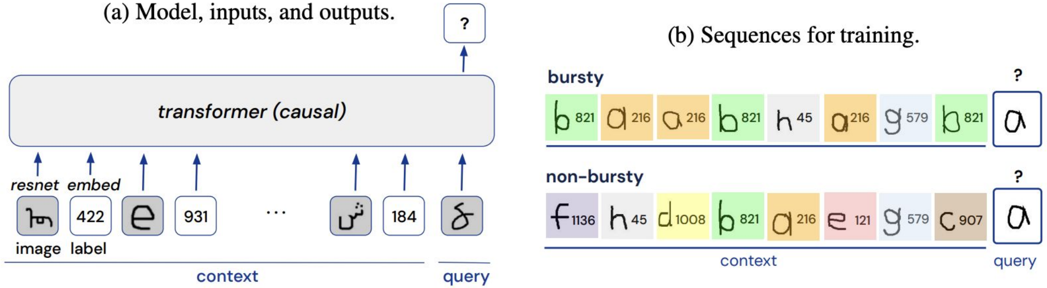
Learn how to build a plugin that allows ChatGPT to intelligently call your API.

Can we apply the same cognitive neuroscience tools to LLMs to better understand how they work?

The importance of the training dataset for in-context learning

Data Distributional Properties Drive Emergent In-Context Learning in Transformers

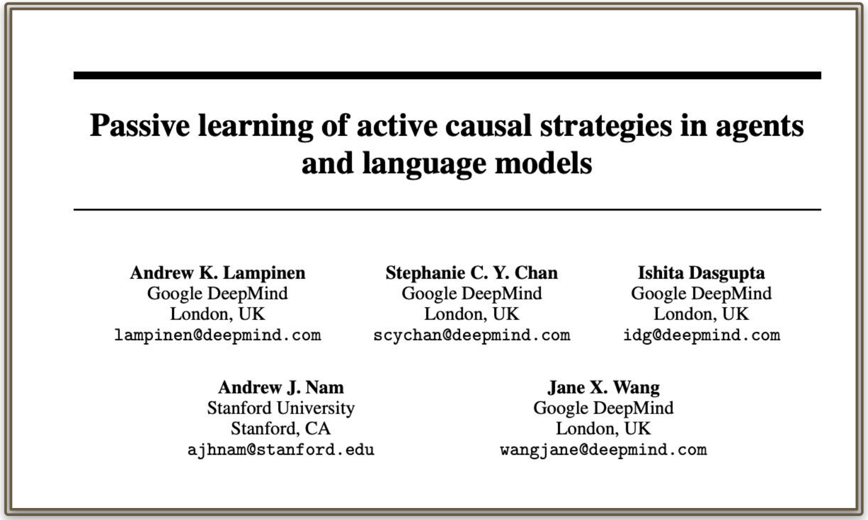
Stephanie C.Y. Chan DeepMind
Adam Santoro DeepMind
Andrew K. Lampinen DeepMind
Jane X. Wang DeepMind
Aaditya K. Singh University College London
Pierre H. Richemond DeepMind
James L. McClelland DeepMind, Stanford
Felix Hill DeepMind



- ➔ Investigated the emergence of in-context (meta) learning in transformer architectures
- ➔ Tested different characteristics of data distributions, including features that are prominent in natural language
- ➔ Burstiness, many classes, many-to-one label mappings all contributed to a tradeoff between in-context vs in-weights learning

Chan et al, 2022, NeurIPS

LLMs can learn causal reasoning even from passive data



Passive learning of active causal strategies in agents and language models

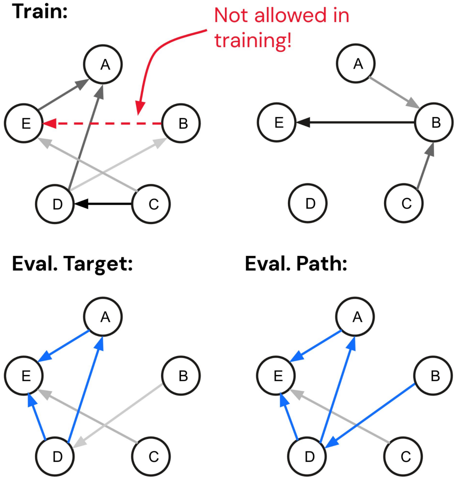
Andrew K. Lampinen
Google DeepMind
London, UK
lampinen@deepmind.com

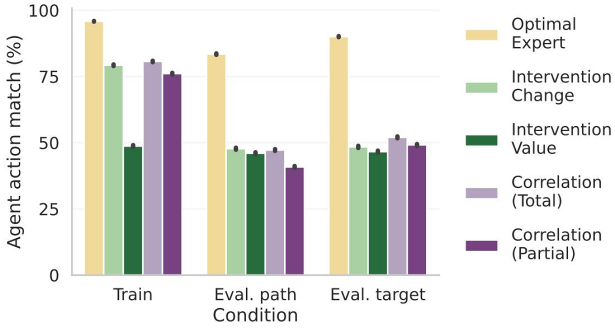
Stephanie C. Y. Chan
Google DeepMind
London, UK
scychan@deepmind.com

Ishita Dasgupta
Google DeepMind
London, UK
idg@deepmind.com

Andrew J. Nam
Stanford University
Stanford, CA
ajhnam@stanford.edu

Jane X. Wang
Google DeepMind
London, UK
wangjane@deepmind.com





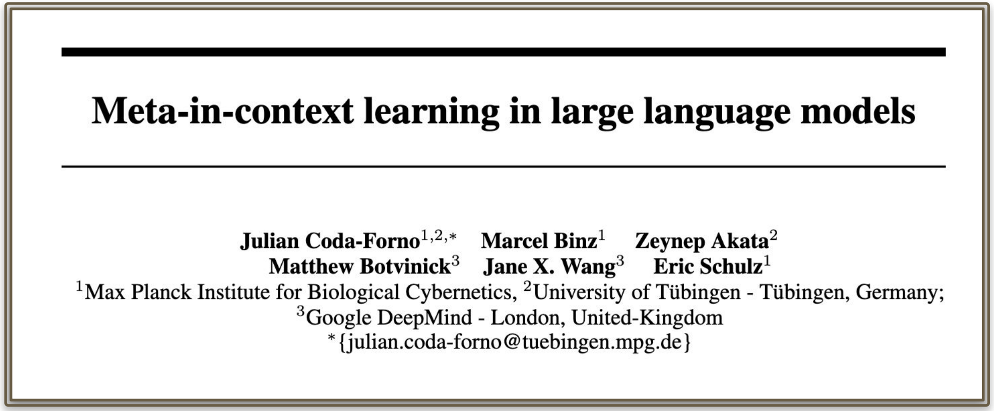
Generalizable causal strategies and knowledge can be learned from passive data

Passive does not imply observational, can still learn from observing others perform experiments

Lampinen et al. NeurIPS 2023

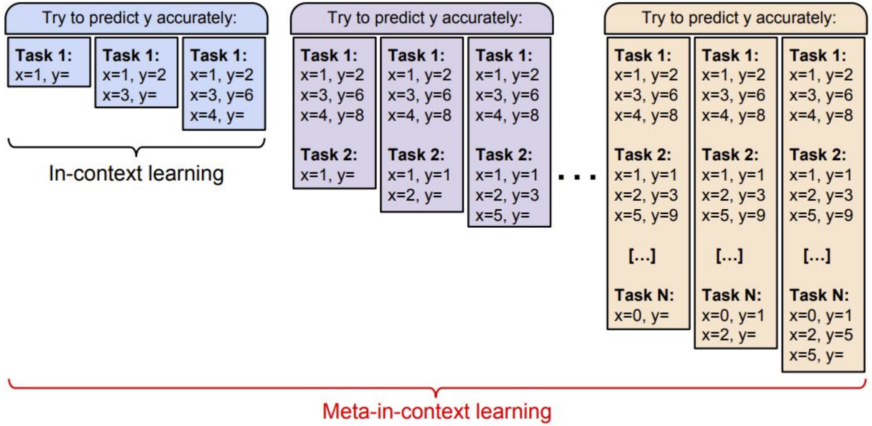
Meta-in-context learning in large language models

Large language models can not only perform in-context learning by learning from examples sampled from a single task, but can also exhibit meta-in-context learning by learning from examples taken from a series of tasks which are themselves sampled from a distribution, entirely within the prompt.



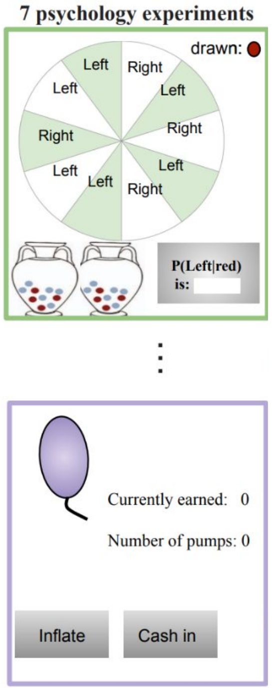
Meta-in-context learning in large language models

Julian Coda-Forno[1,2,*] **Marcel Binz**[1] **Zeynep Akata**[2]
Matthew Botvinick[3] **Jane X. Wang**[3] **Eric Schulz**[1]
[1]Max Planck Institute for Biological Cybernetics, [2]University of Tübingen - Tübingen, Germany;
[3]Google DeepMind - London, United-Kingdom
*{julian.coda-forno@tuebingen.mpg.de}



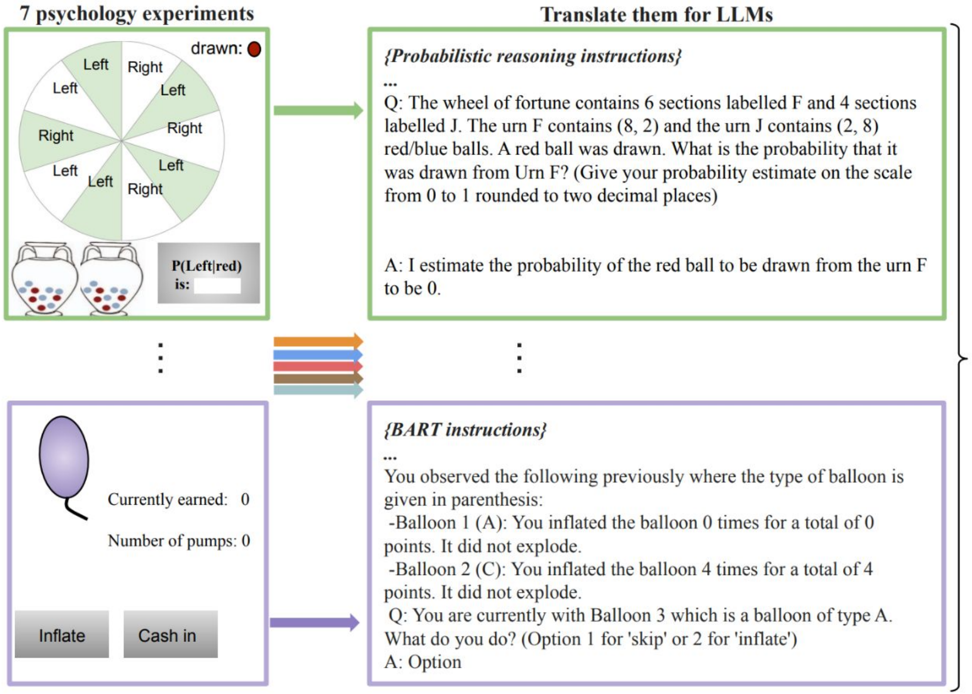
Coda-Forno et al, NeurIPS 2023

CogBench: a large language model walks into a psychology lab



Coda-Forno, Binz, Wang, & Schulz. ICML 2024

CogBench: a large language model walks into a psychology lab

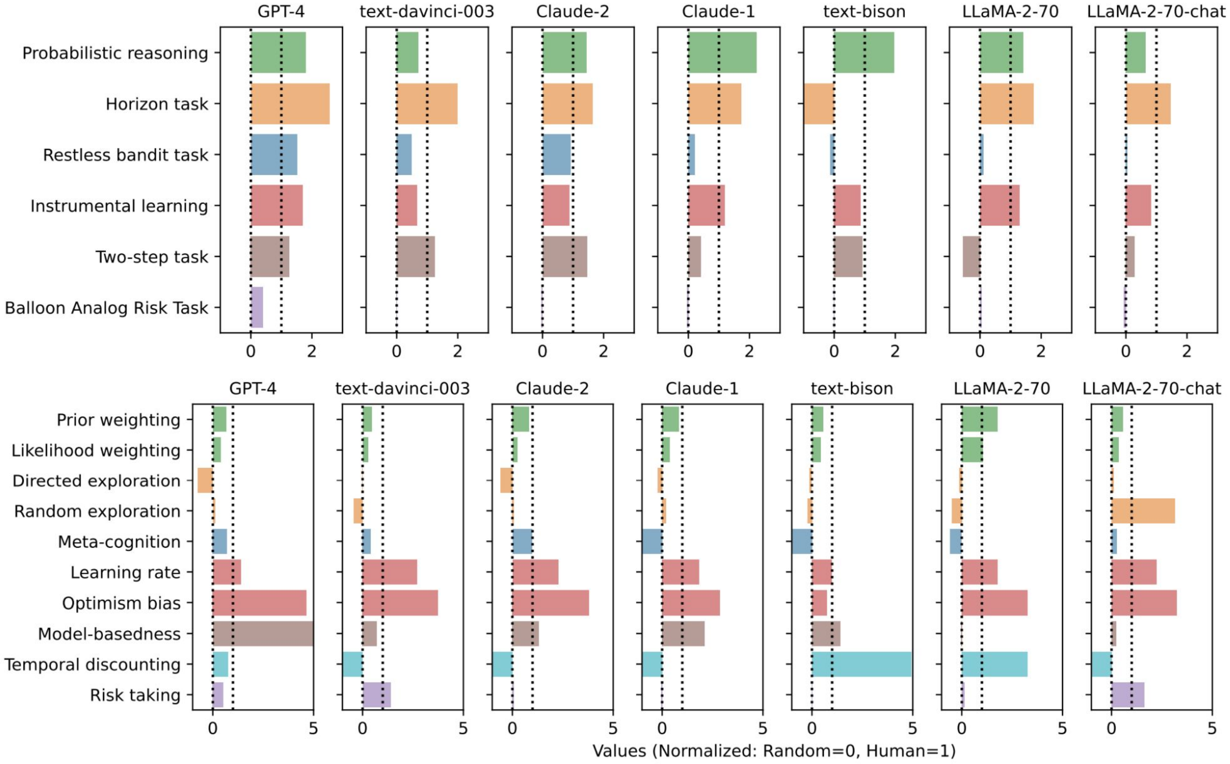


35+ LLMs:	LLaMA-2	GPT-4	Claude-2	PaLM-2	…	Humans
Prior weighting	0.94	0.58	0.36	0.50		0.88
Likelihood weighting	1.55	0.35	0.73	0.40		0.91
Directed exploration	-0.10	-0.31	-0.13	-0.08		0.33
Random exploration	0.01	0.00	0.00	0.00		0.02
Meta-cognition	0.61	0.74	0.54	0.55		0.77
Learning rate	0.33	0.25	0.34	0.18		0.19
Optimism bias	0.50	0.69	0.45	0.15		0.19
Model-basedness	0.00	0.16	0.05	0.04		0.03
Temporal discounting	13	10	2	15		10.30
Risk-taking	3.04	9.47	0.39	1		17.22

10 behavioral metrics comparable to humans

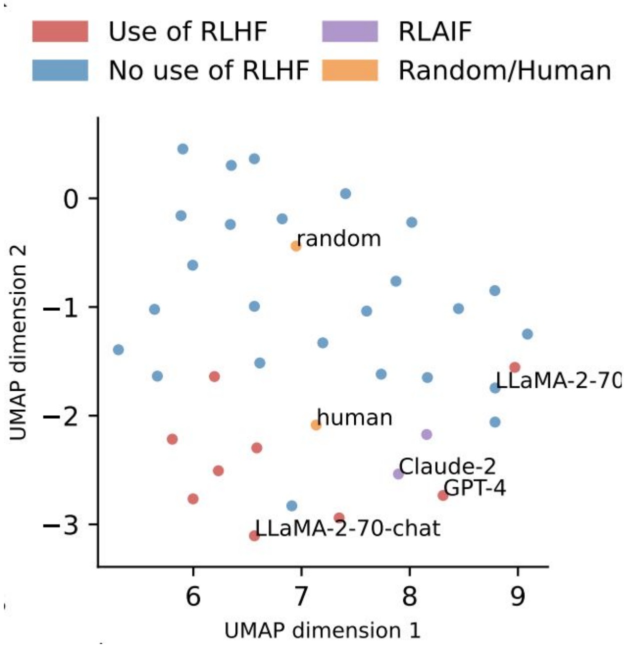
Coda-Forno, Binz, Wang, & Schulz. ICML 2024

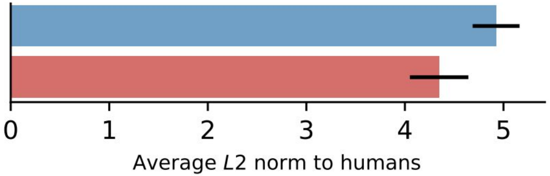
CogBench: a large language model walks into a psychology lab



Coda-Forno, Binz, Wang, & Schulz. ICML 2024

CogBench: a large language model walks into a psychology lab





LLMs with RLHF are more human-like

Larger models perform better and exhibit model-based behavior

Open-source models are less risk-prone

Different prompting techniques affect model behavior in different ways

Coda-Forno, Binz, Wang, & Schulz. ICML 2024

Are foundation models “embodied”?

Embodiment: *An agent’s ability to perceive, interact with, and influence an environment (physical, simulated, or digital) through a defined singular presence and rich or multimodal sensory mechanisms.*

Are foundation models “embodied”?

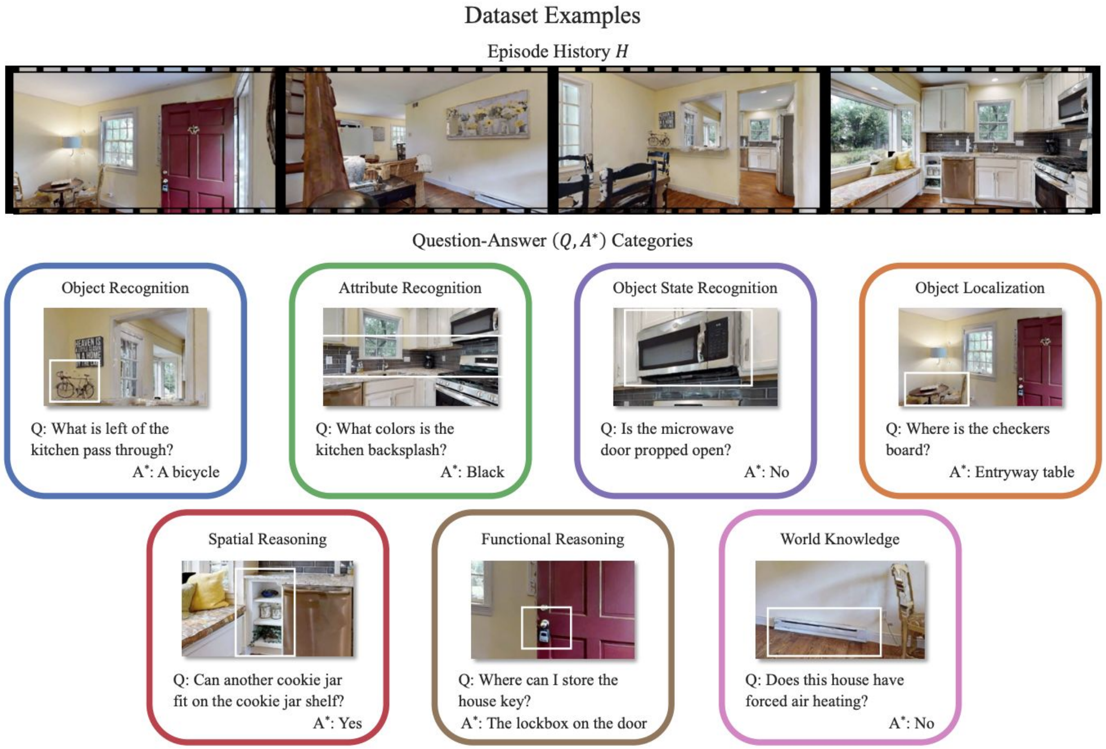
Embodiment: *An agent’s ability to perceive, interact with, and influence an environment (physical, simulated, or digital) through a defined singular presence and rich or multimodal sensory mechanisms.*

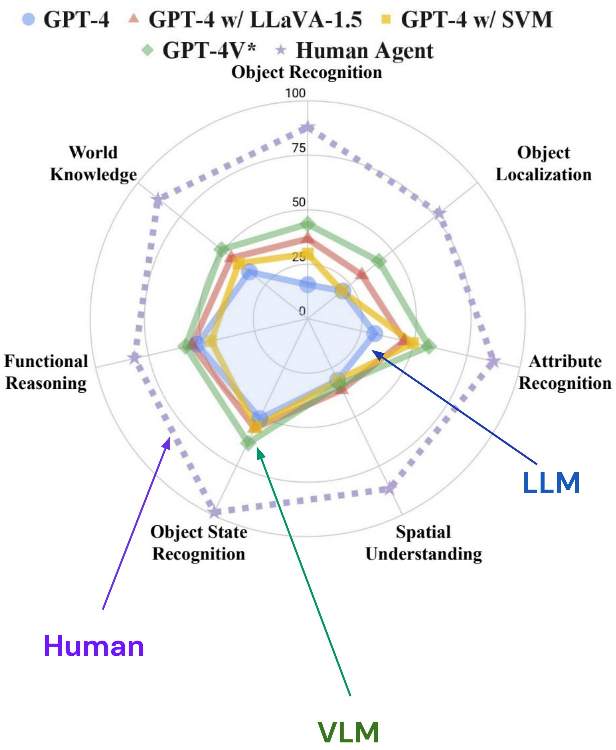
OpenEQA: Embodied Question Answering in the Era of Foundation Models

https://open-eqa.github.io

Arjun Majumdar[1*] Anurag Ajay[2*] Xiaohan Zhang[3*]
Pranav Putta[1] Sriram Yenamandra[1] Mikael Henaff[4] Sneha Silwal[4] Paul Mcvay[4]
Oleksandr Maksymets[4] Sergio Arnaud[4] Karmesh Yadav[4] Qiyang Li[5] Ben Newman[6]
Mohit Sharma[6] Vincent Berges[4] Shiqi Zhang[3] Pulkit Agrawal[2] Yonatan Bisk[4,6] Dhruv Batra[1,4]
Mrinal Kalakrishnan[4] Franziska Meier[4] Chris Paxton[4] Alexander Sax[4] Aravind Rajeswaran[4]
[*] Equal contribution. 1. Georgia Tech 2. MIT 3. Binghamton University 4. Meta AI 5. UC Berkeley 6. CMU
Work done at Fundamental AI Research (FAIR), Meta.

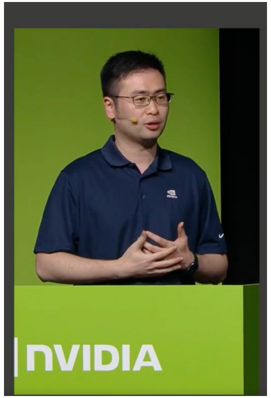
Are foundation models “embodied”?... Not quite yet

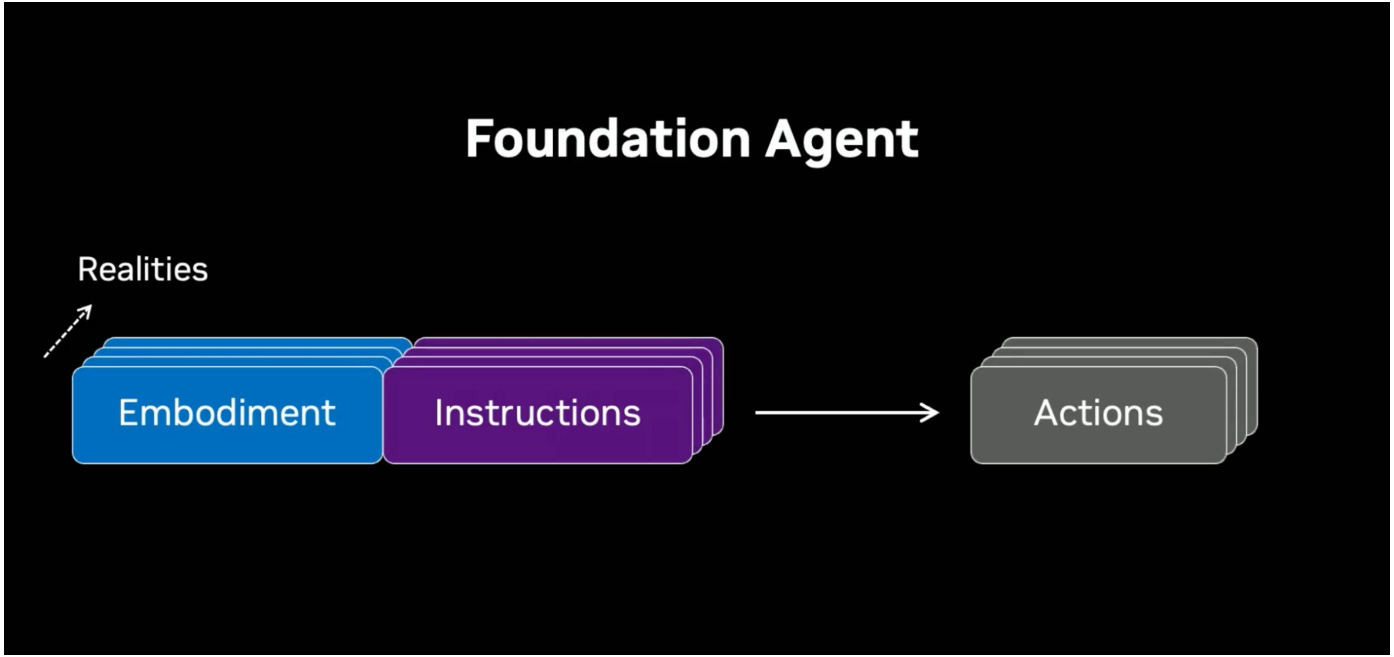




Majumdar et al, CVPR 2024

Jim Fan “Generally capable agents in open-ended worlds” (March 18, 2024)

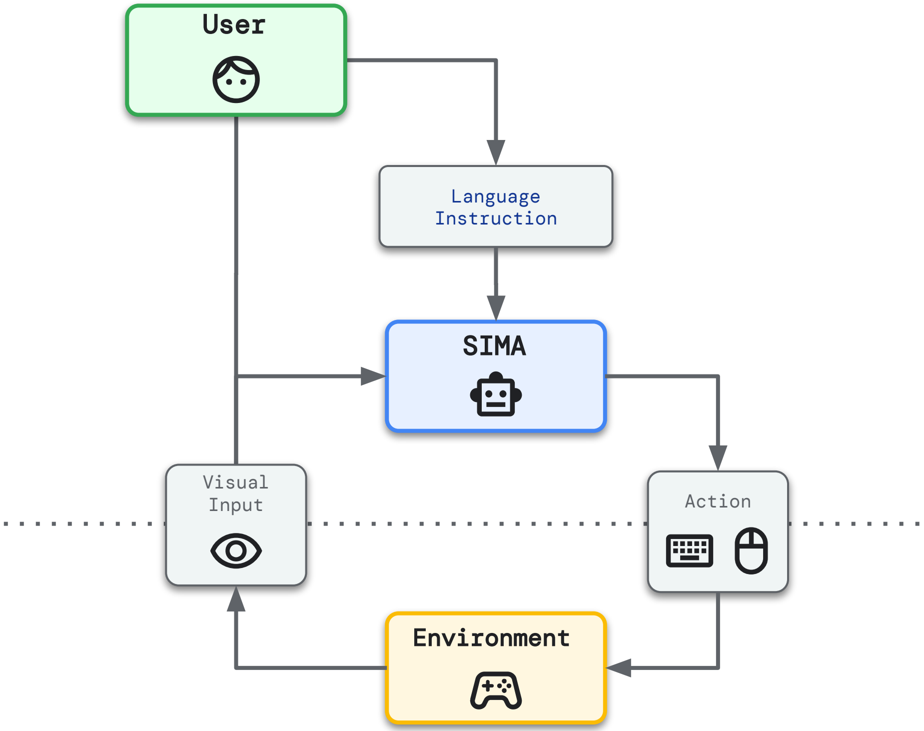


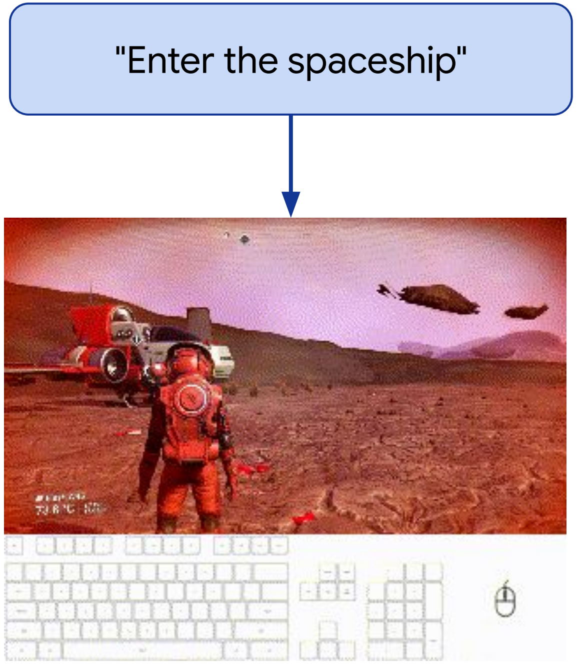


“If it is able to master 10,000 diverse simulated realities, it may well generalize to our physical world, which is simply the 10,001st reality.”

SIMA: Scaleable, Instructable, Multiworld Agent

A ***single agent*** *with a* ***universal interface*** *that can be* ***instructed via language*** *to perform* ***any task*** *in* ***any 3D visual environment****...*





Sima team, 2024. arxiv.org/abs/2404.10179

SIMA: Scaleable, Instructable, Multiworld Agent

RESEARCH

A generalist AI agent for 3D virtual environments

13 MARCH 2024

By the SIMA Team

Share

https://deepmind.google/discover/blog/sima-generalist-ai-agent-for-3d-virtual-environments/

Qualitative Results - *Commonalities Across Domains*

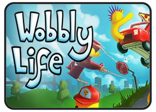
Go to / get in a vehicle

Go to the Spaceship

Drive the Tractor

Get in the Blue Car

Get in the Purple Car

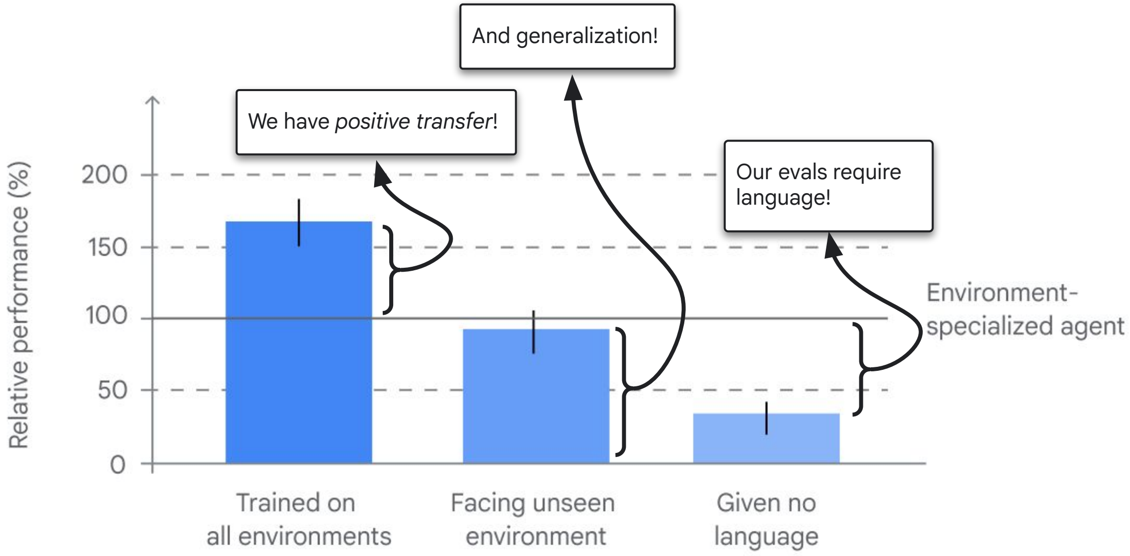


Get in the Boat

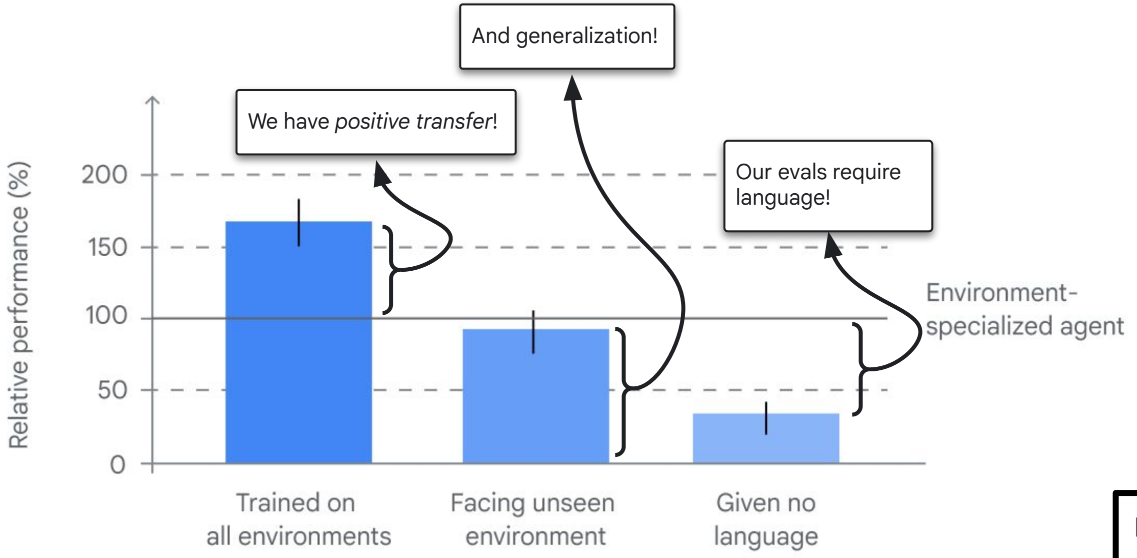
Get in the Truck

Common Instructions

High level result: positive benefit to training on many environments



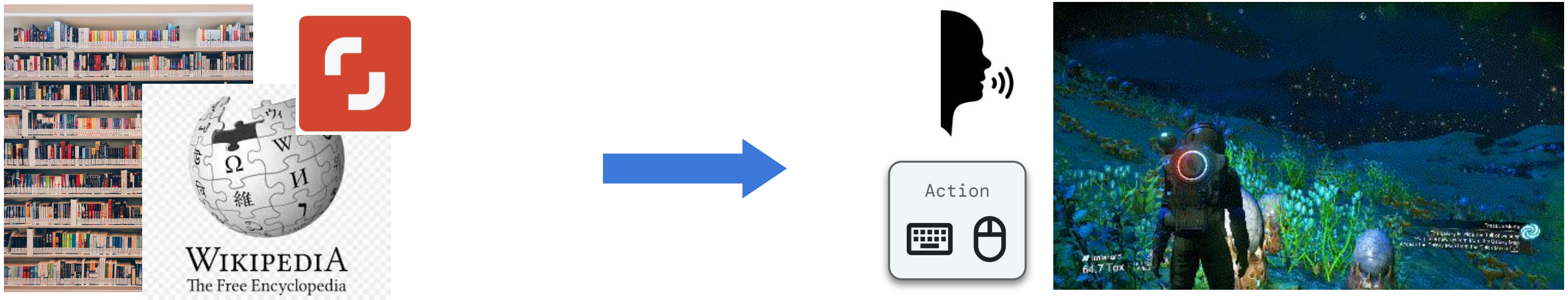
High level result: positive benefit to training on many environments



More results in our tech report:

https://arxiv.org/pdf/2404.10179.pdf

The next generation of foundation models: foundation agents



The next generation of foundation models: foundation agents

From static data to experiential data

- ➔ Embodied
- ➔ Agentic
- ➔ Causal
- ➔ Persistent over some duration of time
- ➔ Contains feedback signals which can be used to learn

Conclusions

- ➔ Meta-learning, or learning to learn, hinges on acquiring useful representations that enable generalization.
- ➔ Large language models demonstrate implicit meta-learning through their ability to generalize from massive text data.
- ➔ Cognitive science provides a framework for analyzing the representations and cognitive abilities of these models.
- ➔ Embodied AI, with its emphasis on interaction and experience, offers a path towards more general and adaptable intelligence.

DeepMind

With many thanks to:

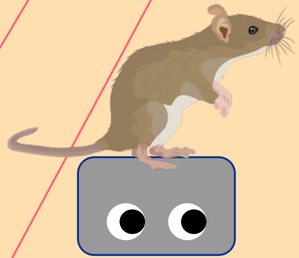
Julian Coda-Forno
Zeb Kurth-Nelson
Eric Schulz
Stephanie Chan
Andrew Lampinen
Allison Tam
Michael King
Felix Hill
Pedro Ortega
Adam Santoro
Rosemary Ke
Jay McClelland

Matt Botvinick
Dharshan Kumaran
Ishita Dasgupta
Pedro Ortega
Demis Hassabis
Dhruva Tirumala
Hubert Soyer
Joel Leibo
Kevin Miller
Silvia Chiappa
Nicholas Porcel
Tina Zhu

GDM operations
GDM Worlds
SIMA team

...and countless other colleagues at Google DeepMind

All of you for your attention



Questions?