Curiosity, Imagination and Information

Applied Machine and Deep Learning
190.015

Vedant Dave

October 2023

Chair of Cyber-Physical-Systems

Outline

- What are our Questions?
- Imagination-augmented agents
- Relevant feature selection in Conscious brain
- Curiosity-driven learning
- What future holds?

What are our Questions?

- **How?**

  How can I represent all the data I see?

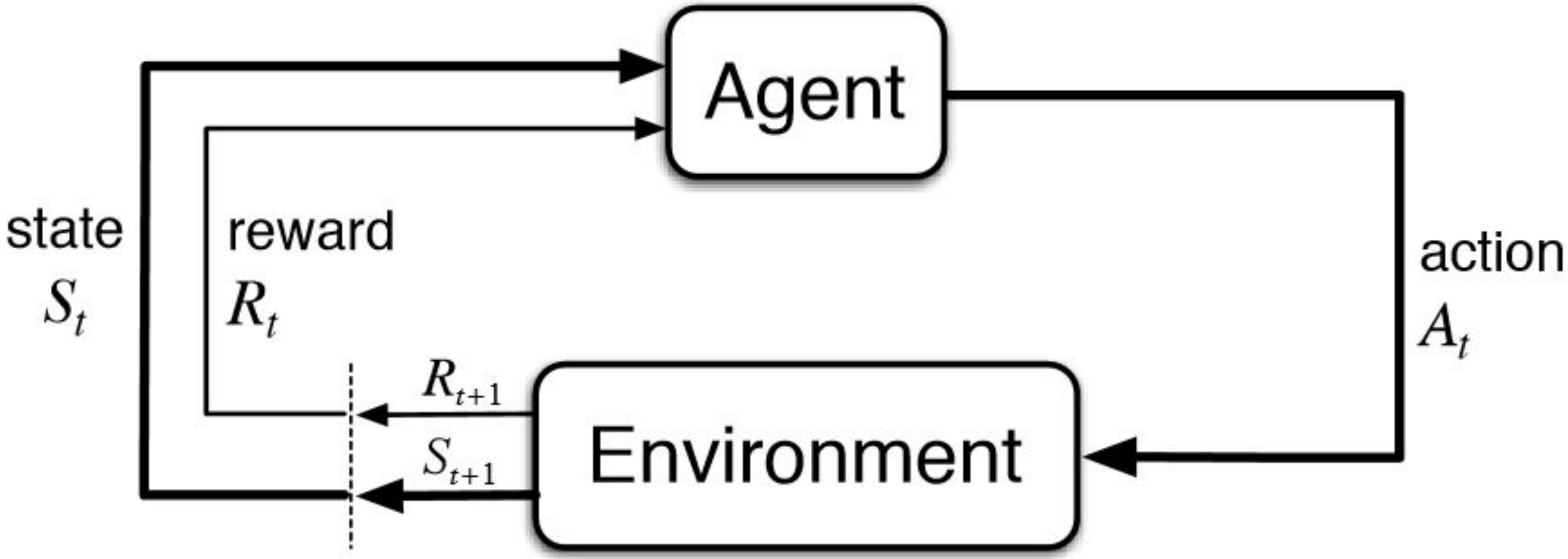
- **What?**

  What should I actually represent from all this mess?

- **Why?**

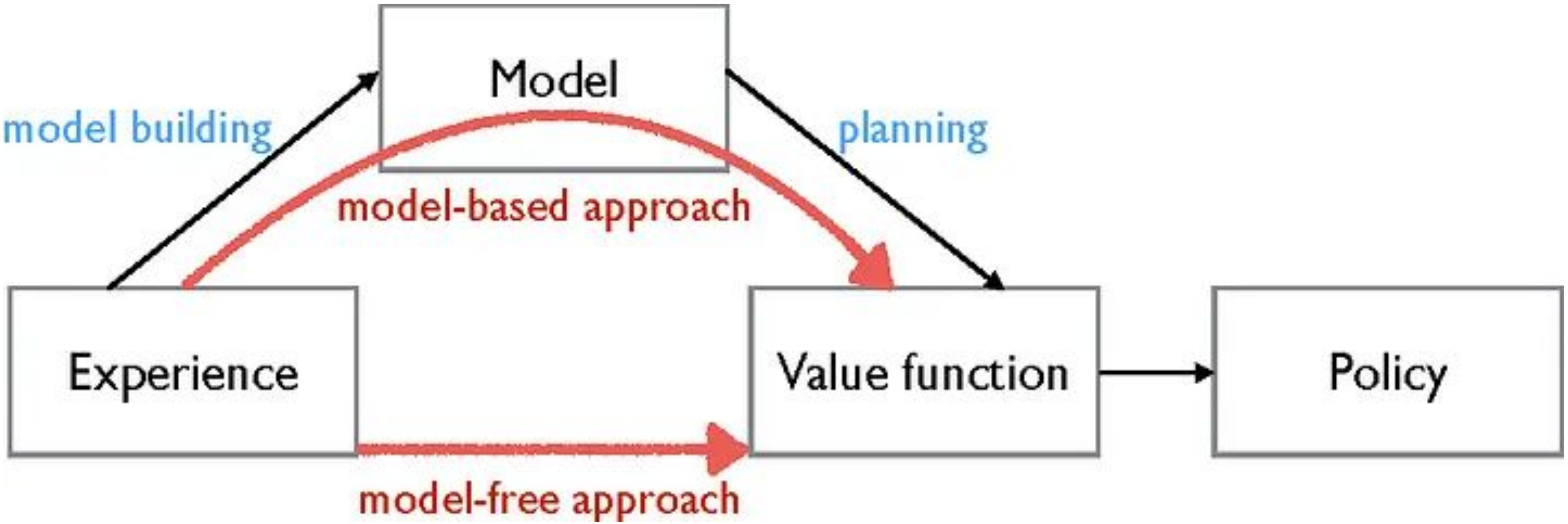
  Why should I do anything at all? Is it worth doing anything if there are no rewards?

Reinforcement Learning



Source: Sutton, R. S., Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. The MIT Press.

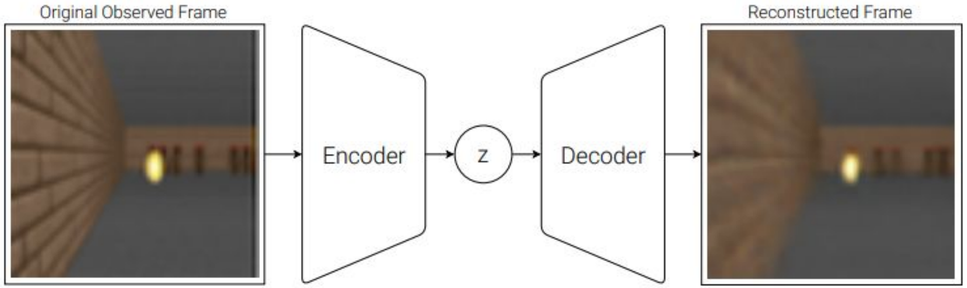
Model-free and Model-based learning



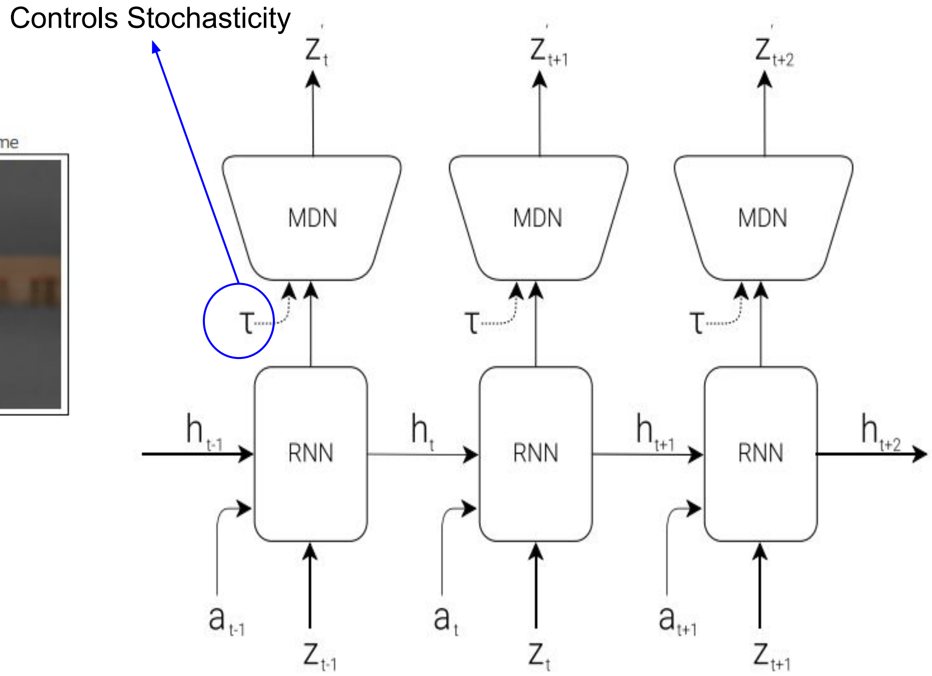
Source: https://medium.com/analytics-vidhya/model-based-offline-reinforcement-learning-morel-f5cd991d9fd5


The Truck Backer-Upper: An Example of Self-Learning in Neural Networks
Derrick Nguyen and Prof. Bernard Widrow
Dept. of Electrical Engineering
Stanford University
Applications of Advances in Nonlinear
Paul J. Werbos, U.S. Depart
An On-Line Algorithm for Dynamic Reinforcement Learning and Planning in Reactive Environments
Jürgen Schmidhuber*
PILCO: A Model-Based and Data-Efficient Approach to Policy Search
Reinforcement Learning in Markovian and Non-Markovian Environments
Jürgen Schmidhuber
Institut für Informatik
Data-Efficient Reinforcement Learning in Continuous-State POMDPs
Rowan McAllister
Carl Rasmussen
From Pixels to Torques: Policy Learning with Deep Dynamical Models
Niklas Wahlström
Thomas B. Schön
Marc Peter Deisenroth
Abstract
Learning deep dynamical models from image pixels
Niklas Wahlström* Thomas B. Schön** Marc Peter Deisenroth***
1 INTRODUCTION

World Models



Variational Autoencoder



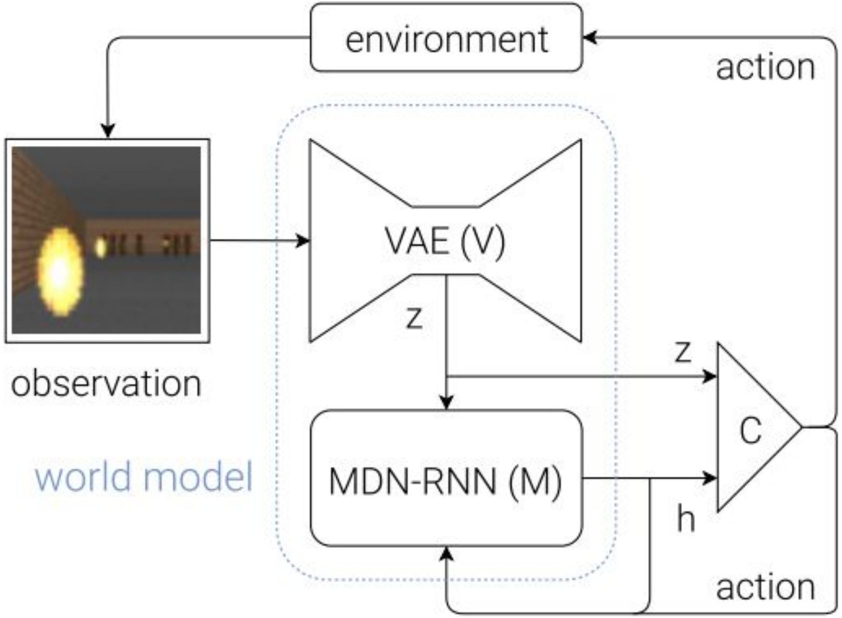
Transition Model

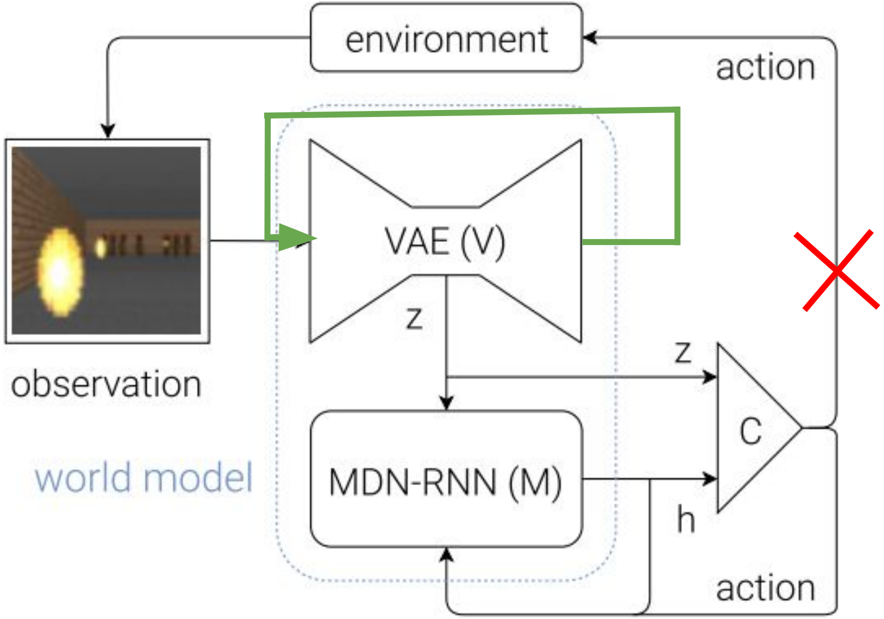
Policy

$$a_t = W_c \left[z_t \; h_t \right] + b_c$$

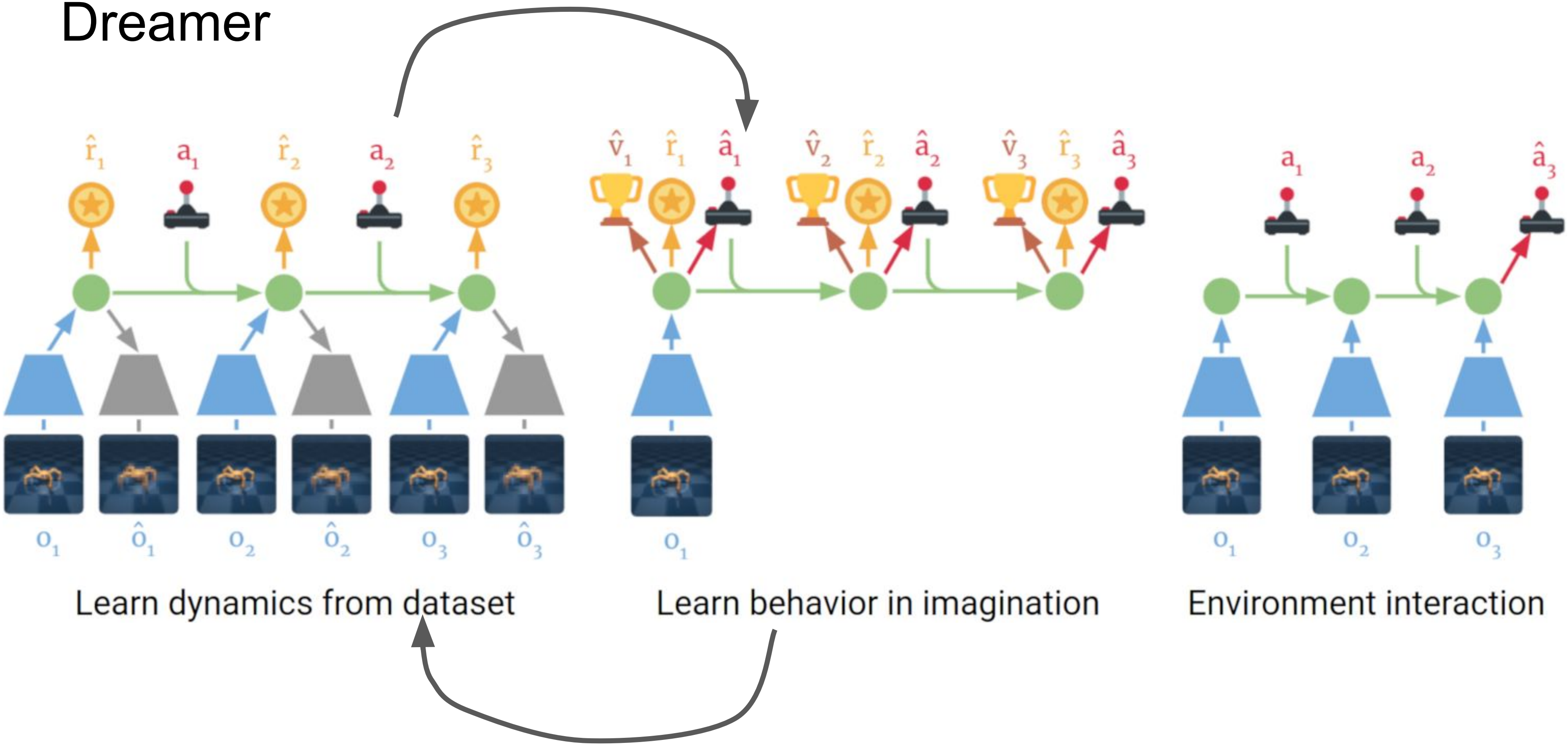
World Models

https://worldmodels.github.io/





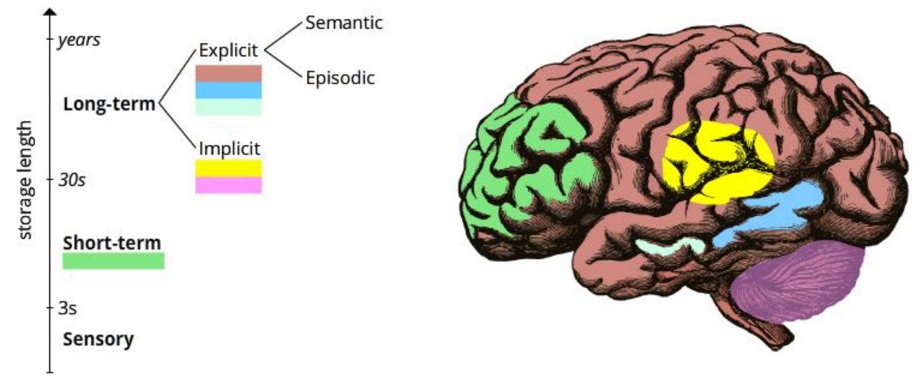
Dreamer


$\hat{r}_1$ a_1 $\hat{r}_2$ a_2 $\hat{r}_3$
o_1 $\hat{o}_1$ o_2 $\hat{o}_2$ o_3 $\hat{o}_3$
$\hat{v}_1$ $\hat{r}_1$ $\hat{a}_1$ $\hat{v}_2$ $\hat{r}_2$ $\hat{a}_2$ $\hat{v}_3$ $\hat{r}_3$ $\hat{a}_3$
o_1
a_1 a_2 $\hat{a}_3$
o_1 o_2 o_3
Learn dynamics from dataset
Learn behavior in imagination
Environment interaction

Dreamer

Dreamer-v2

Relevant Feature selection in Conscious Brain



Activations may not be exact

https://gohighbrow.com/the-brain-and-memory/

- **Short-term Memory:** Holds seven, plus minus two, pieces of information [Miller]

  Lets Play a Game

  Not very long-lasting

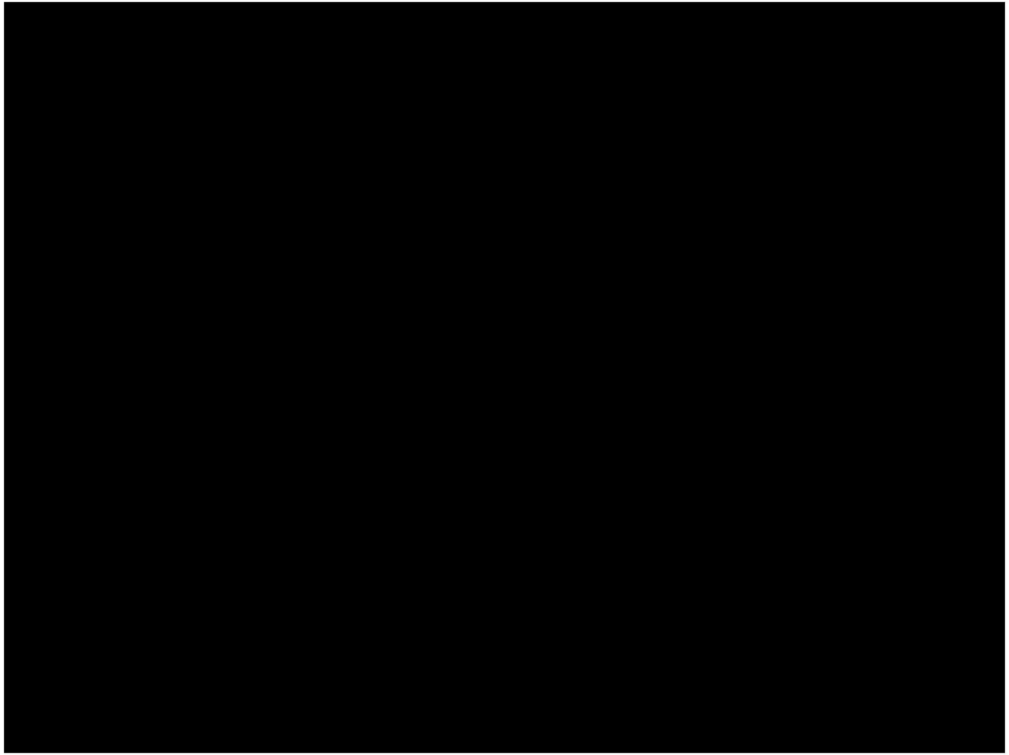
  You need attention

  Numbers are much better and Relevancy matters [Jacobs]
- Do we need so much information to solve a task? (Selective Attention Test)
- How to select relevant features from this noisy world? What is even meant by "relevant"?
- Can we imagine noise-free worlds from noisy observations?

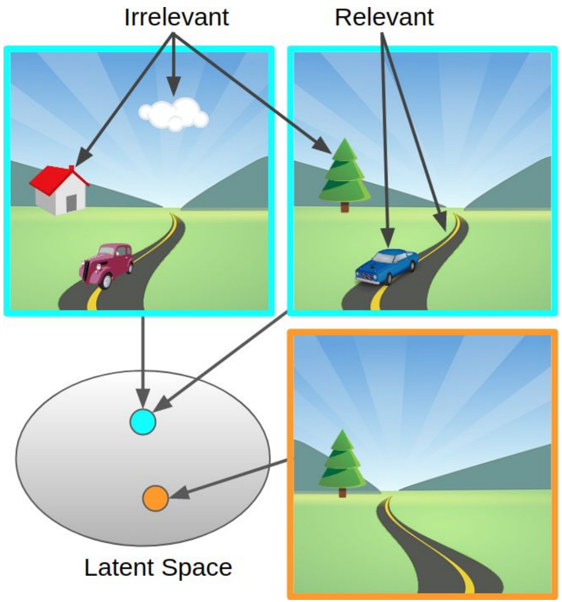
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. https://doi.org/10.1037/h0043158

Joseph Jacobs, Experiments on "Prehension", Mind, Volume os-12, Issue 45, 1 January 1887, Pages 75–79.

- **One of my inspiration**



Research in relevancy prediction



Deep Bisimulation for Control

Bisimulation Metric: Measures similarity between similar states

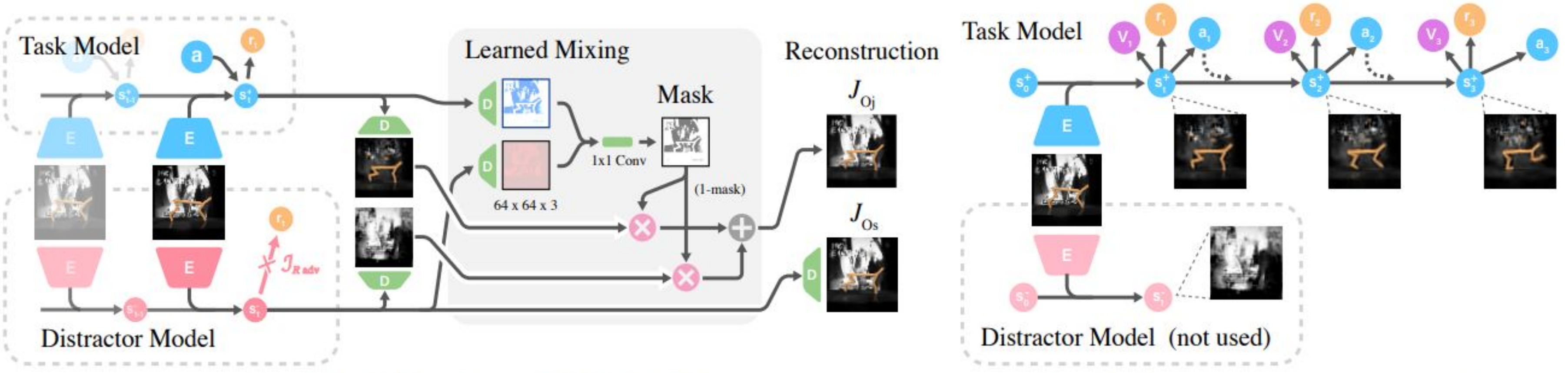
Zhang, Amy, et al. "Learning invariant representations for reinforcement learning without reconstruction." arXiv preprint arXiv:2006.10742 (2020).

Some Assumptions

- Used same video for test and train
- No RGB background
- Used Ground Plane

Research in relevancy prediction

Learning Task Informed Abstractions



(a) Learning **Task Informed World Models**

(b) **Policy Learning** only unrolls in $\mathcal{S}^{+}$.

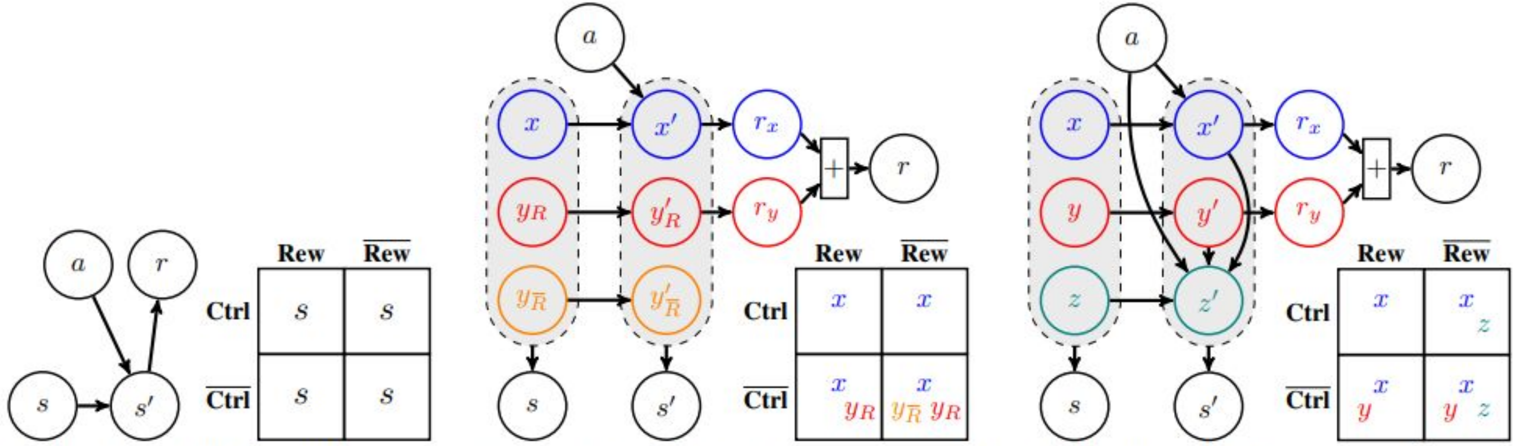
Fu, Xiang, et al. "Learning task informed abstractions." International Conference on Machine Learning. PMLR, 2021.

Learning Task Informed Abstractions

Fu, Xiang, et al. "Learning task informed abstractions." International Conference on Machine Learning. PMLR, 2021.

Research in relevancy prediction

Denoised MDPs



(a) Transition without useful structure. s may contain any type of information.

(b) Transition that factorizes out uncontrollable information in $y_{\bar{R}}$ and y_R.

(c) Transition that factorizes out uncontrollable y and reward-irrelevant z.

Tongzhou Wang, Simon S. Du, Antonio Torralba, Phillip Isola, Amy Zhang, and Yuandong Tian. Denoised mdps: Learning world models better than the world itself. In International Conference on Machine Learning. PMLR, 2022.

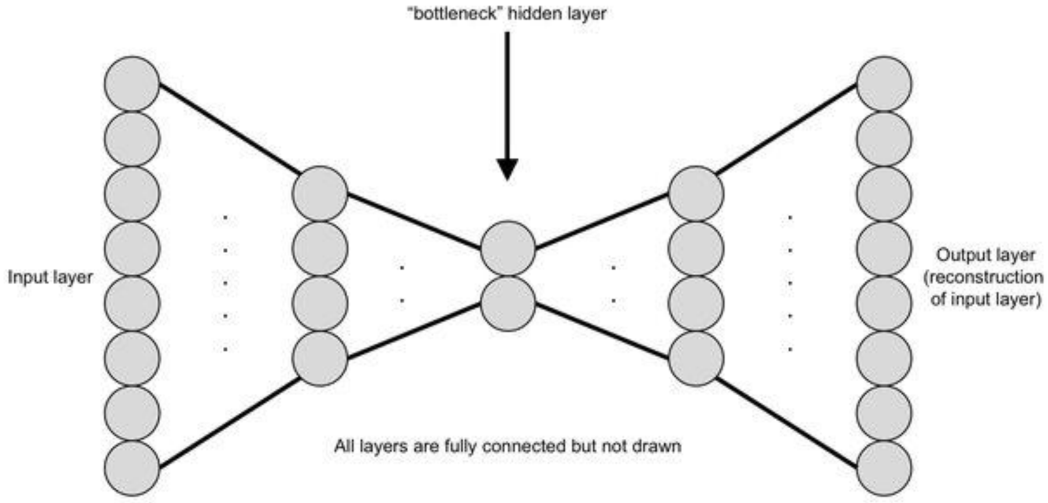
Denoised MDPs

Tongzhou Wang, Simon S. Du, Antonio Torralba, Phillip Isola, Amy Zhang, and Yuandong Tian. Denoised mdps: Learning world models better than the world itself. In International Conference on Machine Learning. PMLR, 2022.

Our Research

Denoised Predictive Imagination

- Information Bottleneck Principle

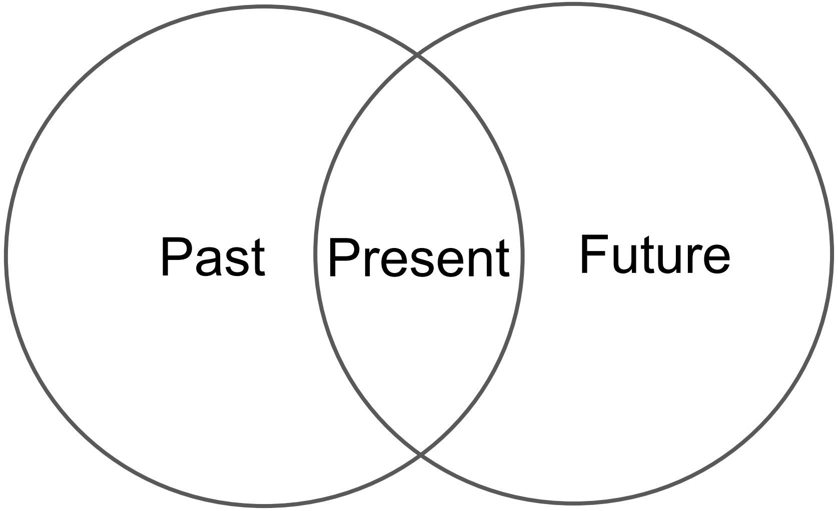


Tishby, Naftali, Fernando C. Pereira, and William Bialek. "The information bottleneck method." *arXiv preprint physics/0004057* (2000).

Our Research

Denoised Predictive Imagination

- Predictive Information

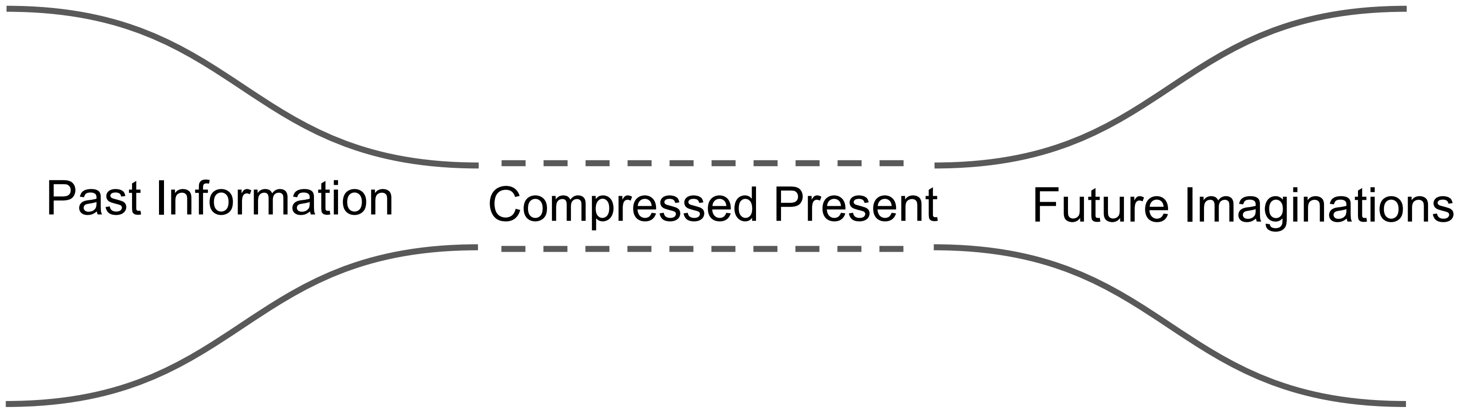


Bialek, William, and Naftali Tishby. "Predictive information." *arXiv preprint cond-mat/9902341* (1999).

Our Research

Denoised Predictive Imagination

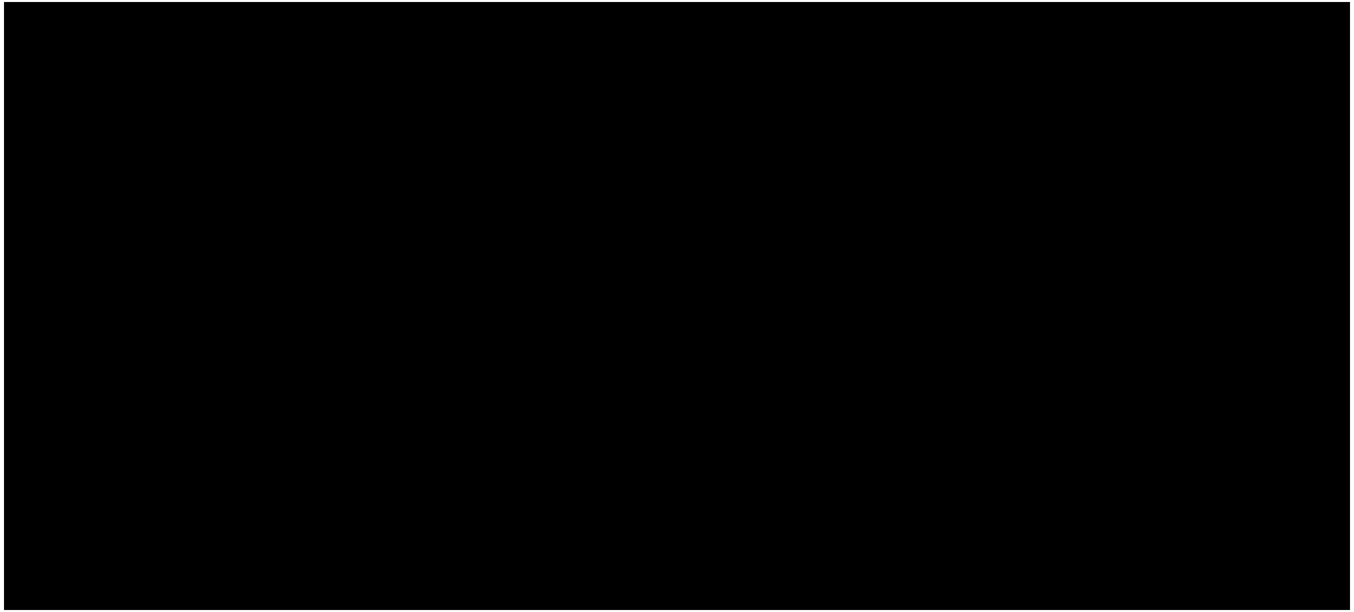
- Information Bottleneck + Predictive Information + Soft Actor-Critic (Reinforcement Learning)



Submitted to International Conference of Learning Representations (ICLR) 2024

Our Research

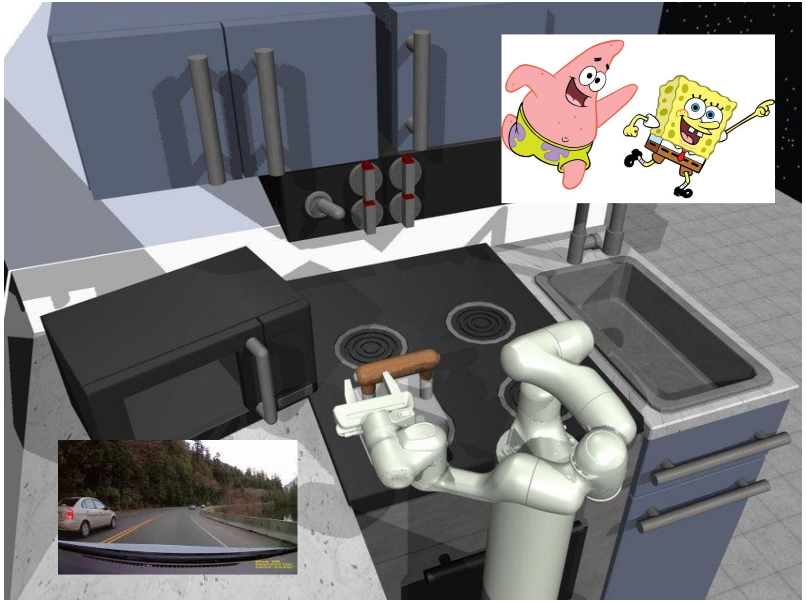
Denoised Predictive Imagination



Submitted to International Conference of Learning Representations (ICLR) 2024

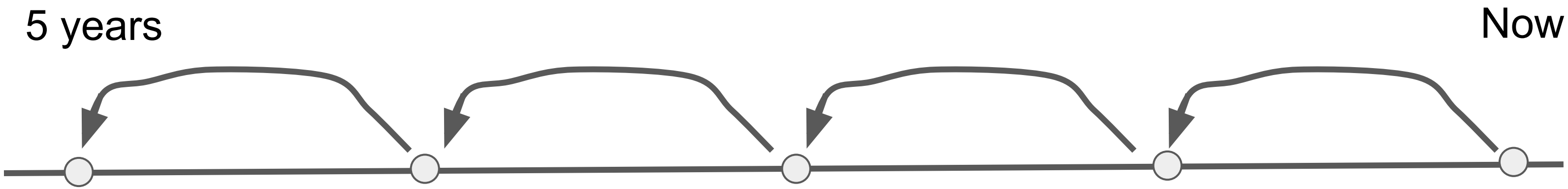
New Ideas

Why not Attention?



Done by ChatGPT-4

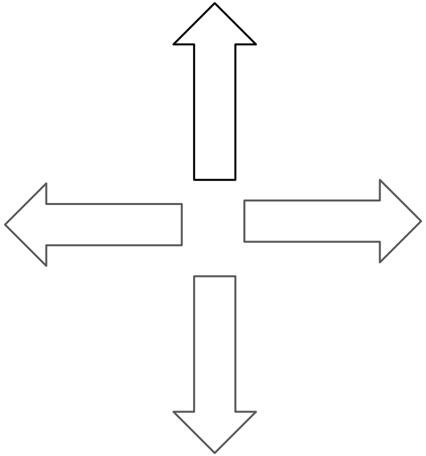
Curiosity/Intrinsic Motivation



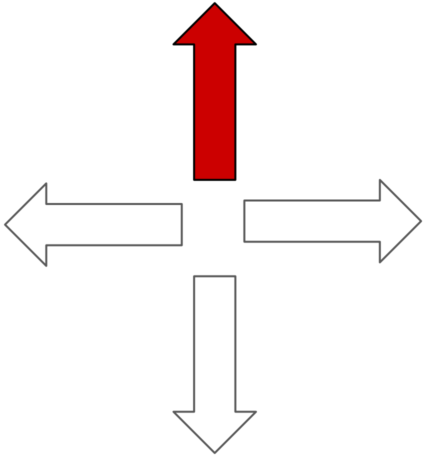
Curiosity/Intrinsic Motivation

Silvia, P. J. (2012). Curiosity and motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 157–166). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399820.013.0010

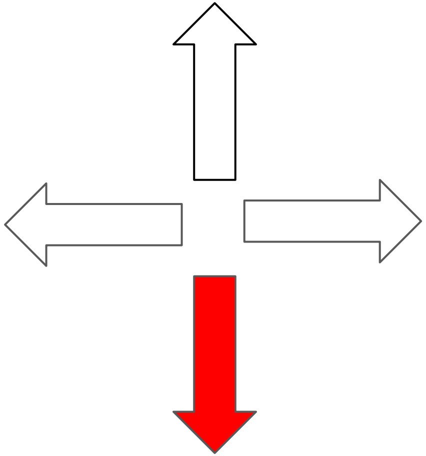
How to formulate Curiosity?



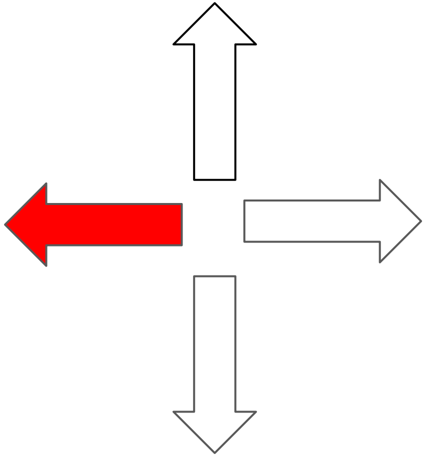
Source: https://supermario-game.com/de

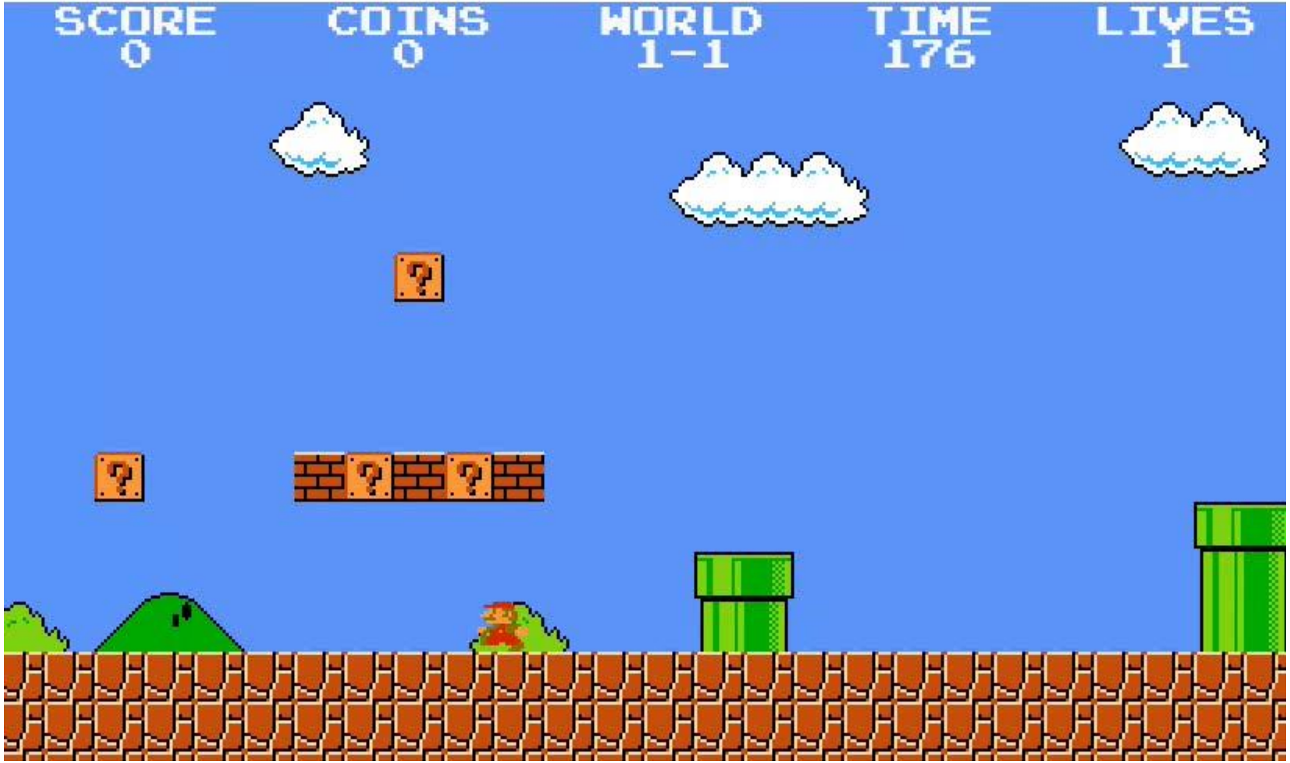



SCORE
0
COINS
0
WORLD
1-1
TIME
190
LIVES
1

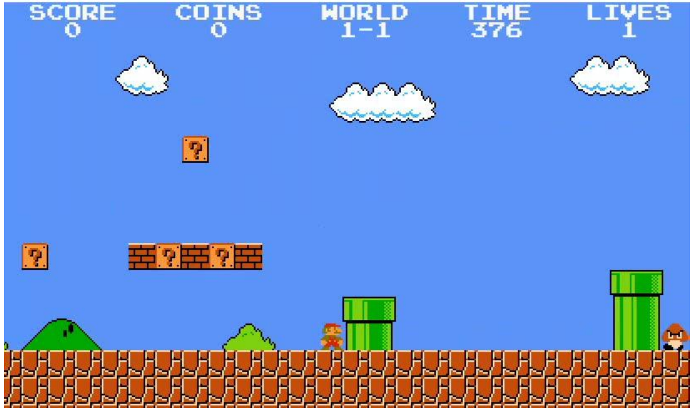



SCORE
0
COINS
0
WORLD
1-1
TIME
184
LIVES
1



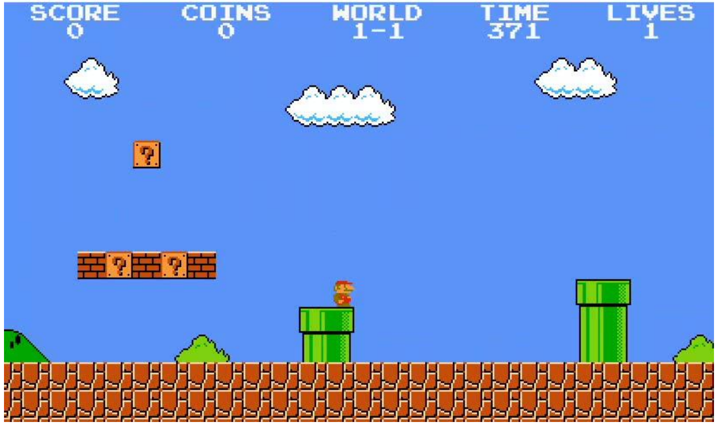

SCORE
0
COINS
0
WORLD
1-1
TIME
176
LIVES
1

Curiosity = Prediction Error



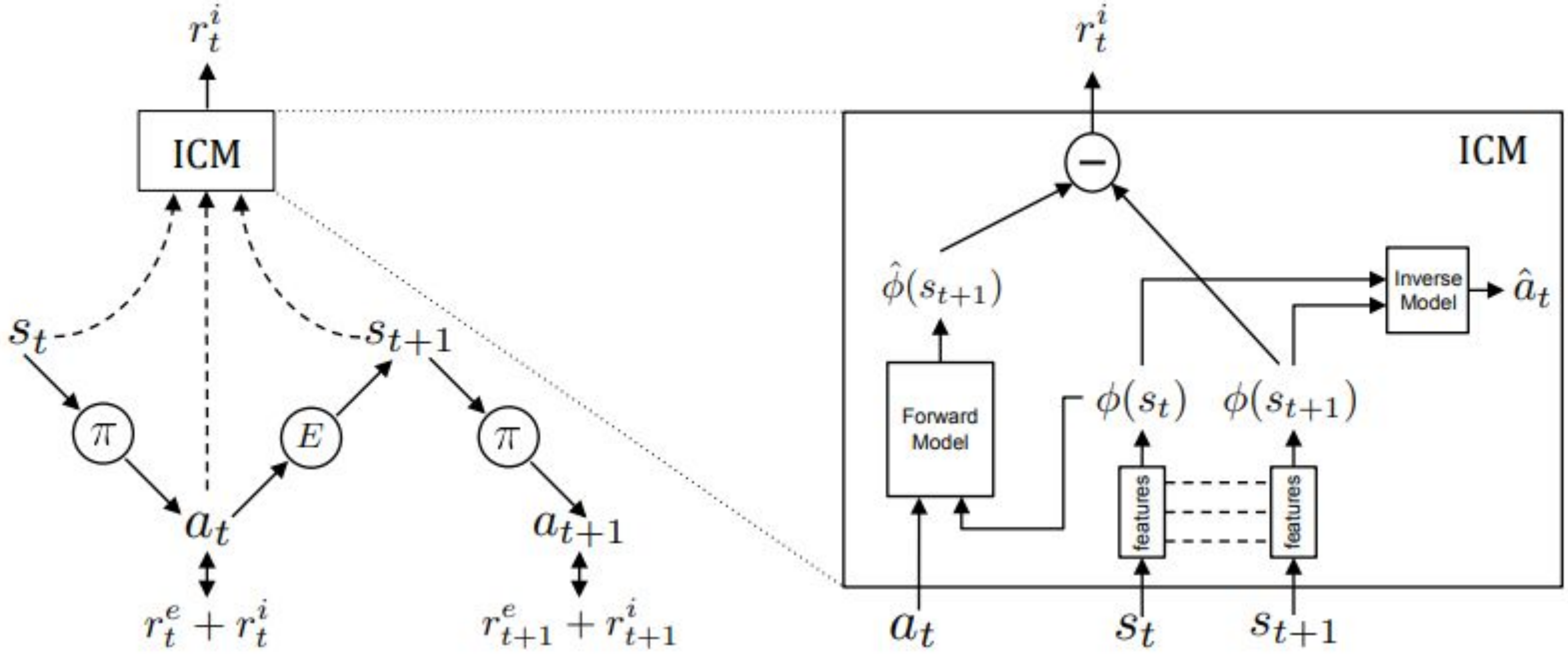
Expected Behaviour

Actual Behaviour



Pathak, Deepak, et al. "Curiosity-driven exploration by self-supervised prediction." *International conference on machine learning*. PMLR, 2017.

Curiosity/Intrinsic Motivation

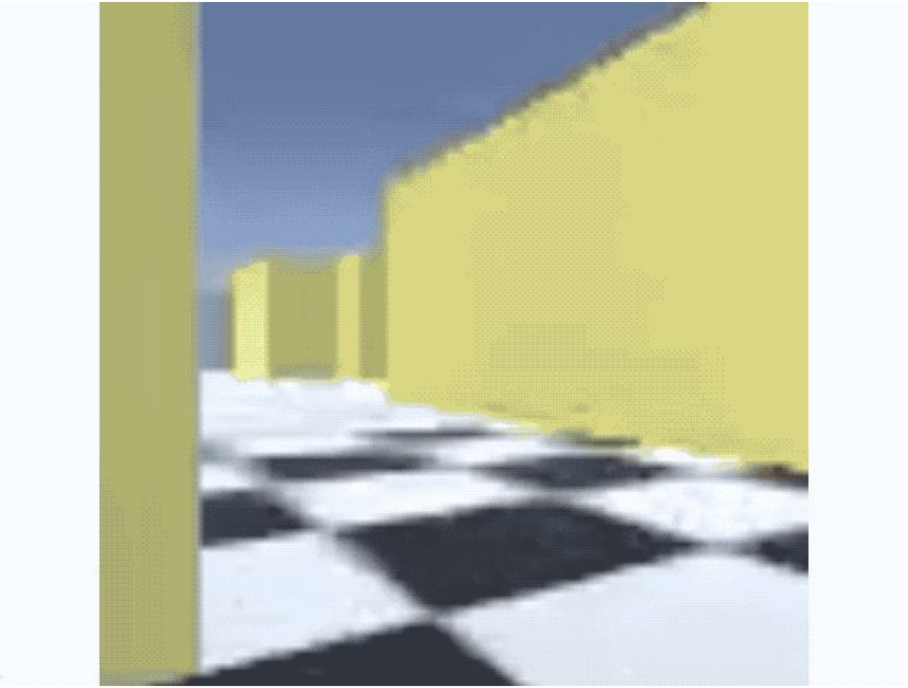


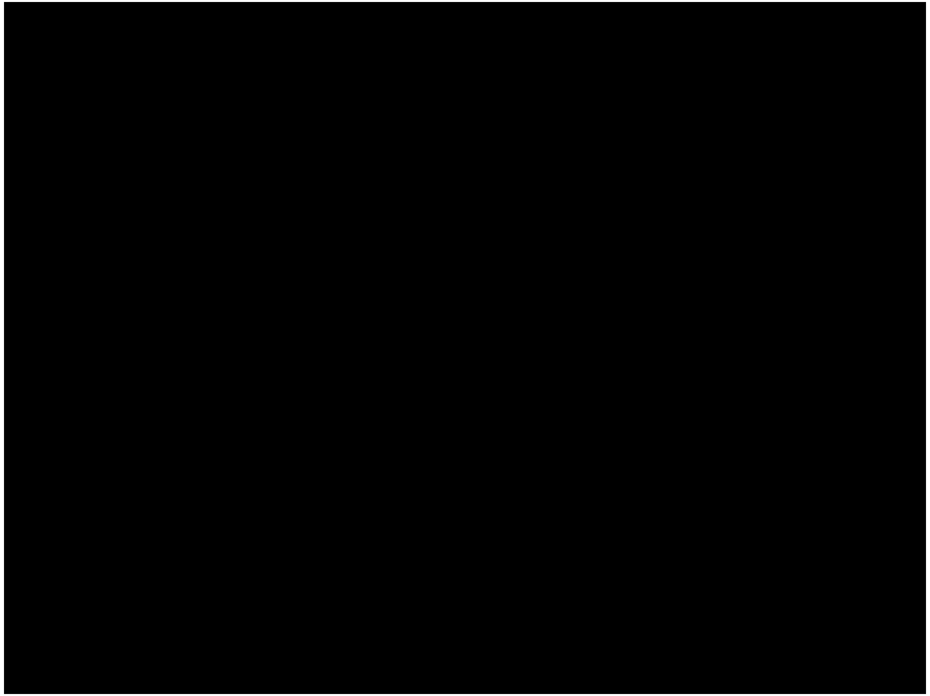
Pathak, Deepak, et al. "Curiosity-driven exploration by self-supervised prediction." *International conference on machine learning*. PMLR, 2017.

Curiosity/Intrinsic Motivation

Pathak, Deepak, et al. "Curiosity-driven exploration by self-supervised prediction." *International conference on machine learning*. PMLR, 2017.

Curiosity/Intrinsic Motivation





Burda, Yuri, et al. "Large-scale study of curiosity-driven learning." *arXiv preprint arXiv:1808.04355* (2018).

Pong, Vitchyr H., et al. "Skew-fit: State-covering self-supervised reinforcement learning." *arXiv preprint arXiv:1903.03698* (2019).

What future holds?

- Better representation in Reinforcement Learning
- Learning with Causal Models
- Long-term Planning
- New Formulations on Curiosity

Credits

From Deep Learning of Disentangled Representations to Higher-level Cognition (https://www.youtube.com/watch?v=Yr1mOzC93xs&t=1529s)

Curiosity: Autonomy Talks - Deepak Pathak: Robotics In The Wild: Continual Improvement by Watching & Practicing (https://www.youtube.com/watch?v=mBkbnQoknr8)