POLITECNICO MILANO 1863 | DIPARTIMENTO DI MECCANICA

11th International Conference on Machine Learning, Optimization and Data Science

Deep Active Inference Agents
for Delayed and Long-Horizon Environments

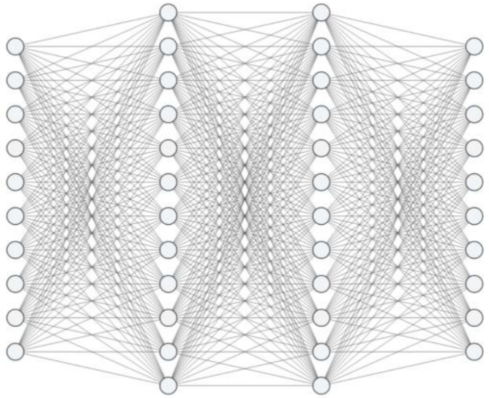
Yavar Taheri Yeganeh [1], Mohsen Jafari [2], Andrea Matta [1]

yavar.taheri@polimi.it, jafari@soe.rutgers.edu, andrea.matta@polimi.it

[1] Politecnico di Milano, [2] Rutgers University

Neuroscience in Machine Learning and AI

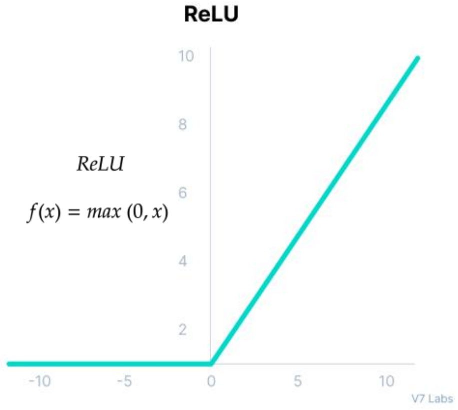
Arterial Neural Networks



Modern Machine Learning
Universal Function Approximators

Inspired by
Biological Neural Networks

ReLU Activation Function



Deep Learning

Inspired by neural firings in the brain "*Neurons either fire or do not fire (binary activation)*"

Convolutional Neural Networks

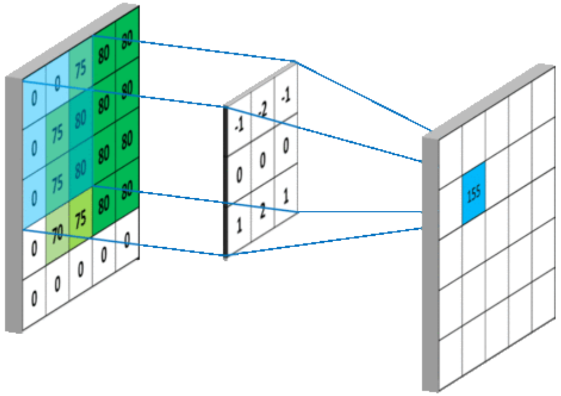
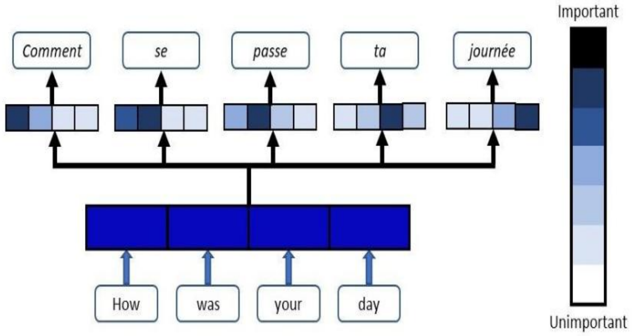


Image Processing
Computer Vision

Inspired by the visual processing in the brain (the hierarchical structure of the visual cortex)

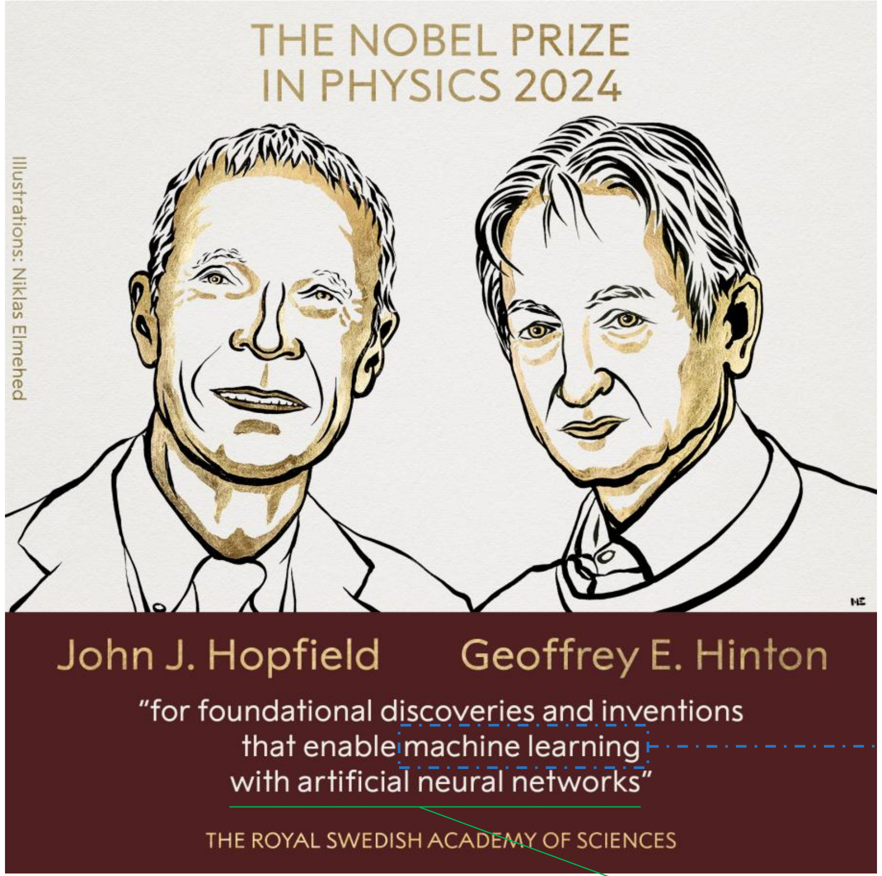
Attention Mechanism



Transformers
LLMs (e..g, ChatGPT)

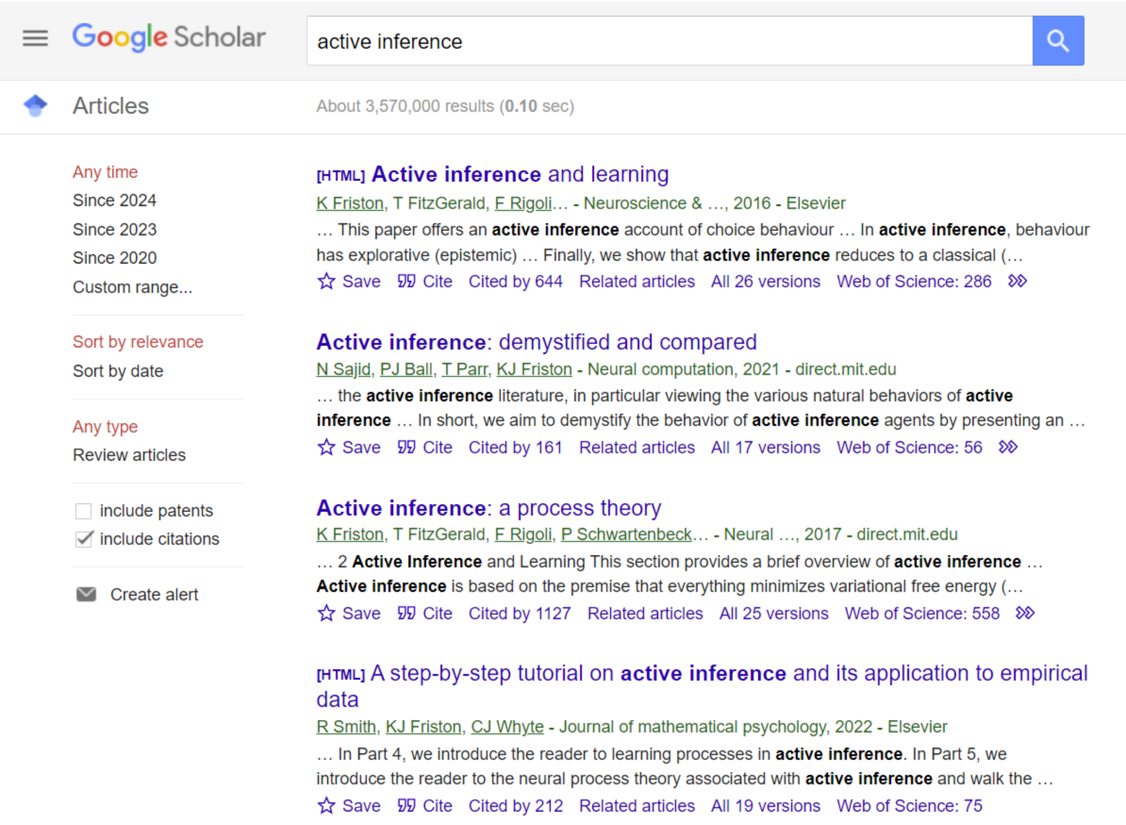
Connections to modulations of neural activities in the brain (selectively focus on specific stimuli)

Neuroscience in Machine Learning and AI



Notably impacted by concepts from Neuroscience (and with Mathematics and Physics)

Neuroscience: Active Inference Trend


Google Scholar
active inference
Articles
About 3,570,000 results (**0.10** sec)

Any time
Since 2024
Since 2023
Since 2020
Custom range...

Sort by relevance
Sort by date

Any type
Review articles

include patents
include citations

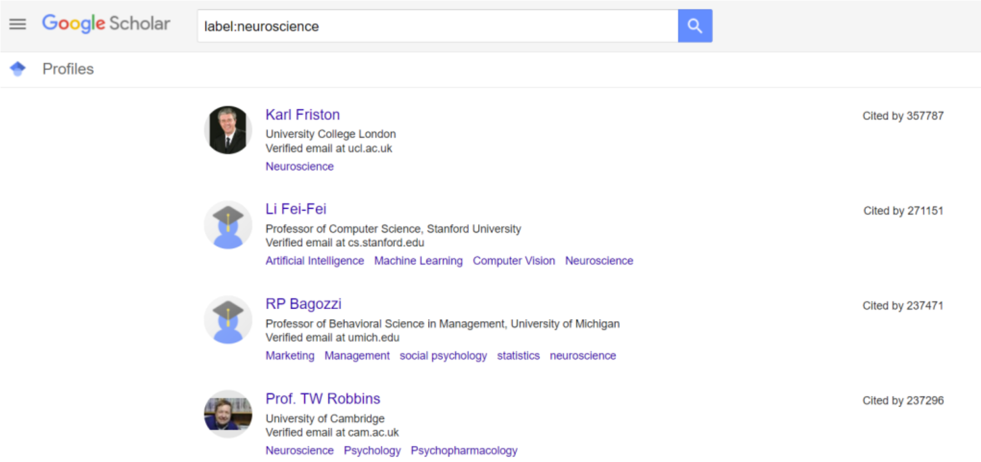
Create alert

[HTML] **Active inference** and learning
K Friston, T FitzGerald, F Rigoli… - Neuroscience & …, 2016 - Elsevier
… This paper offers an **active inference** account of choice behaviour … In **active inference**, behaviour has explorative (epistemic) … Finally, we show that **active inference** reduces to a classical (…
Save Cite Cited by 644 Related articles All 26 versions Web of Science: 286

Active inference: demystified and compared
N Sajid, PJ Ball, T Parr, KJ Friston - Neural computation, 2021 - direct.mit.edu
… the **active inference** literature, in particular viewing the various natural behaviors of **active inference** … In short, we aim to demystify the behavior of **active inference** agents by presenting an …
Save Cite Cited by 161 Related articles All 17 versions Web of Science: 56

Active inference: a process theory
K Friston, T FitzGerald, F Rigoli, P Schwartenbeck… - Neural …, 2017 - direct.mit.edu
… 2 **Active Inference** and Learning This section provides a brief overview of **active inference** … **Active inference** is based on the premise that everything minimizes variational free energy (…
Save Cite Cited by 1127 Related articles All 25 versions Web of Science: 558

[HTML] A step-by-step tutorial on **active inference** and its application to empirical data
R Smith, KJ Friston, CJ Whyte - Journal of mathematical psychology, 2022 - Elsevier
… In Part 4, we introduce the reader to learning processes in **active inference**. In Part 5, we introduce the reader to the neural process theory associated with **active inference** and walk the …
Save Cite Cited by 212 Related articles All 19 versions Web of Science: 75


Google Scholar
label:neuroscience
Profiles

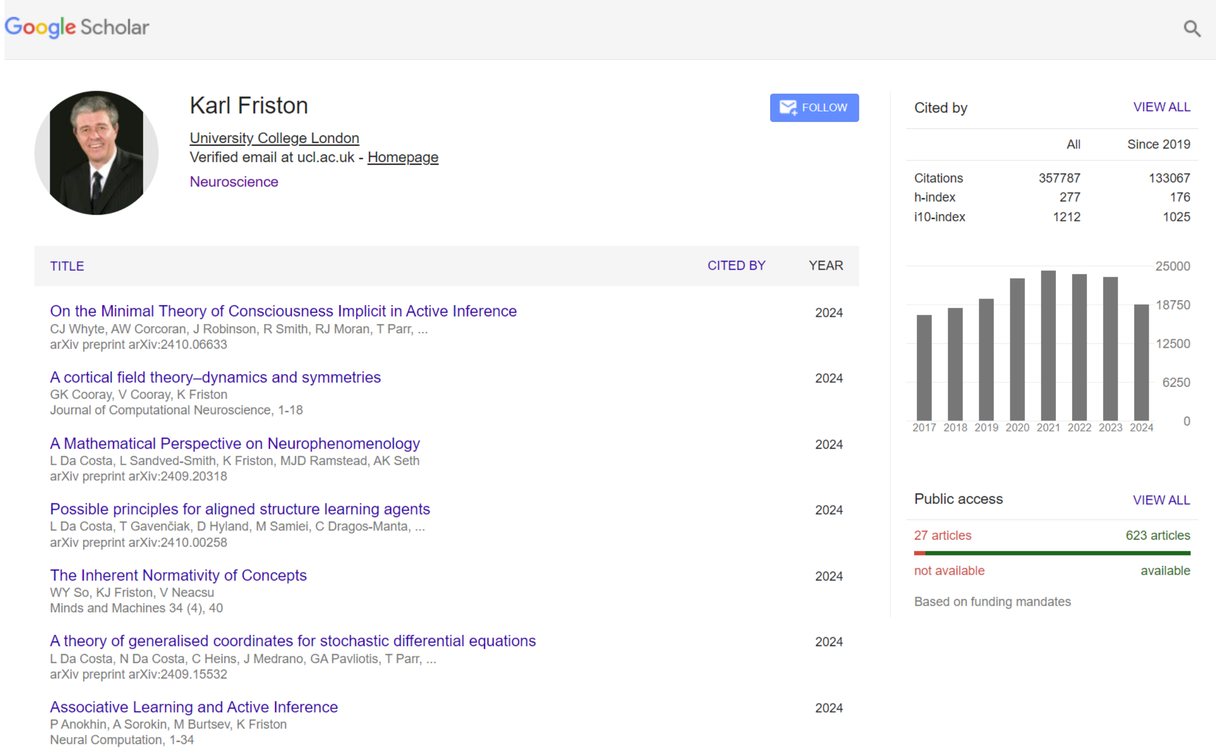
Karl Friston — Cited by 357787
University College London
Verified email at ucl.ac.uk
Neuroscience

Li Fei-Fei — Cited by 271151
Professor of Computer Science, Stanford University
Verified email at cs.stanford.edu
Artificial Intelligence Machine Learning Computer Vision Neuroscience

RP Bagozzi — Cited by 237471
Professor of Behavioral Science in Management, University of Michigan
Verified email at umich.edu
Marketing Management social psychology statistics neuroscience

Prof. TW Robbins — Cited by 237296
University of Cambridge
Verified email at cam.ac.uk
Neuroscience Psychology Psychopharmacology

Karl Friston


Google Scholar

Karl Friston
University College London
Verified email at ucl.ac.uk - Homepage
Neuroscience
FOLLOW

TITLE	CITED BY	YEAR
On the Minimal Theory of Consciousness Implicit in Active Inference CJ Whyte, AW Corcoran, J Robinson, R Smith, RJ Moran, T Parr, ... arXiv preprint arXiv:2410.06633		2024
A cortical field theory–dynamics and symmetries GK Cooray, V Cooray, K Friston Journal of Computational Neuroscience, 1-18		2024
A Mathematical Perspective on Neurophenomenology L Da Costa, L Sandved-Smith, K Friston, MJD Ramstead, AK Seth arXiv preprint arXiv:2409.20318		2024
Possible principles for aligned structure learning agents L Da Costa, T Gavenčiak, D Hyland, M Samiei, C Dragos-Manta, ... arXiv preprint arXiv:2410.00258		2024
The Inherent Normativity of Concepts WY So, KJ Friston, V Neacsu Minds and Machines 34 (4), 40		2024
A theory of generalised coordinates for stochastic differential equations L Da Costa, N Da Costa, C Heins, J Medrano, GA Pavliotis, T Parr, ... arXiv preprint arXiv:2409.15532		2024
Associative Learning and Active Inference P Anokhin, A Sorokin, M Burtsev, K Friston Neural Computation, 1-34		2024

Cited by VIEW ALL

	All	Since 2019
Citations	357787	133067
h-index	277	176
i10-index	1212	1025

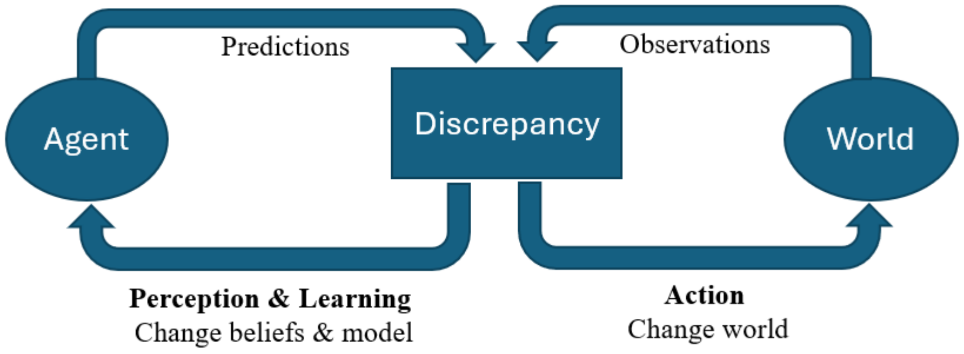
Public access VIEW ALL
27 articles 623 articles
not available available
Based on funding mandates

Active Inference: Concept

Theory of Biological Brains

Based on Predictive Coding: Both are grounded on Generative Model and Predictions

Perception and Learning are driven by Prediction Errors.



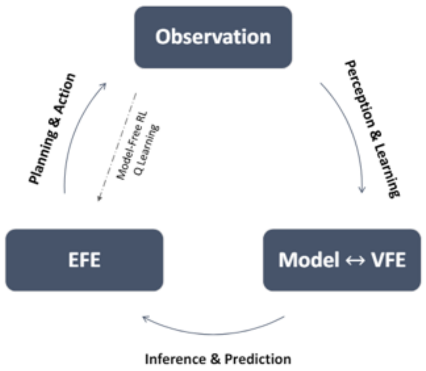
Observations are influenced by actions: $\boldsymbol{Observation\ (action)}$

Predictive Coding → Active Inference

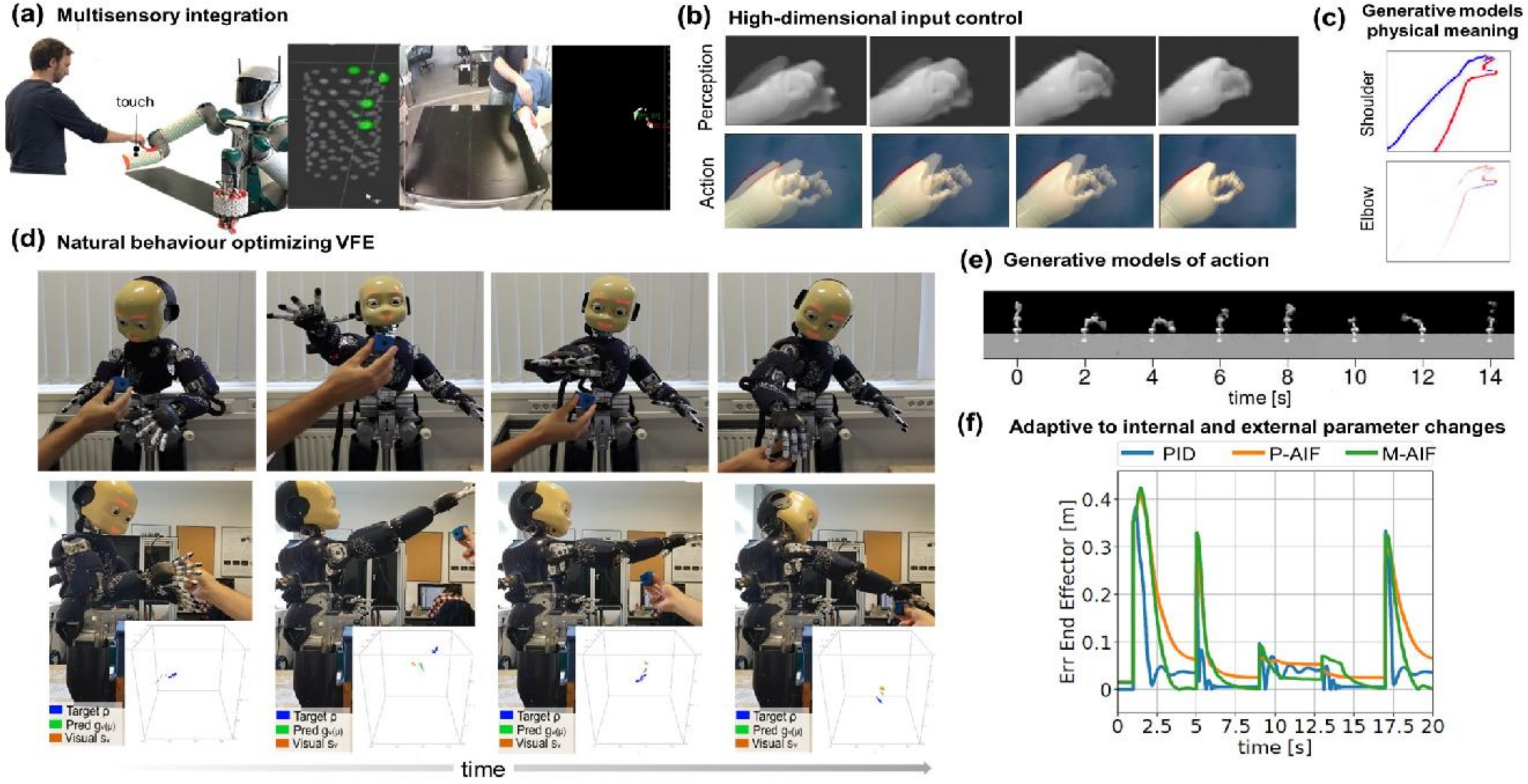
Free Energy Principle: Minimizing surprise $(-\ln p(observation|model)\ \text{or}\ Free\ Energy)$

Active Inference unifies perception, learning, and action:

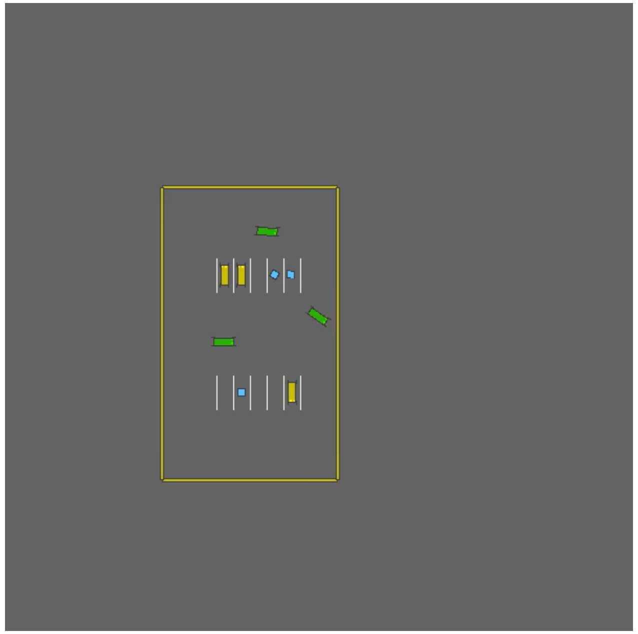
- Variational Free Energy: Perception & Learning Model
- Expected Free Energy: Action



Active Inference: Robotics



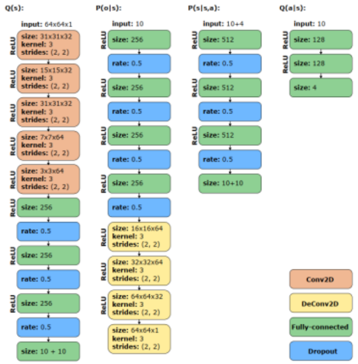
Robot Control [2]



Navigation of Autonomous Vehicles [3]

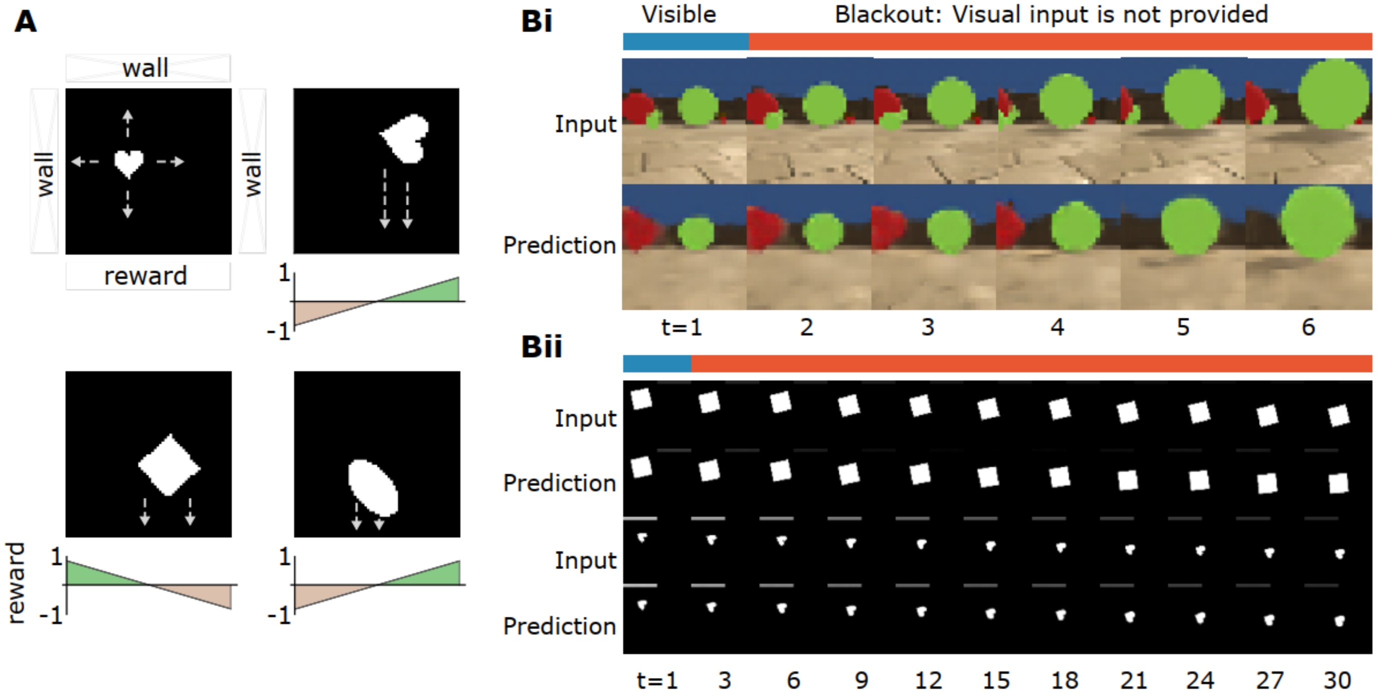
[2] Lanillos et al., 2021
[3] Huang et al., 2024

Active Inference: Against Model-Free (DAIMC [4])

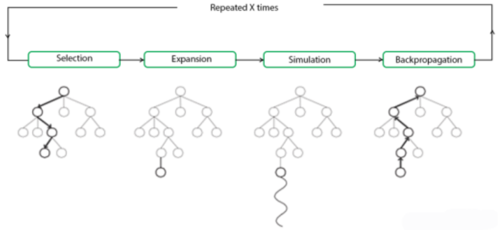


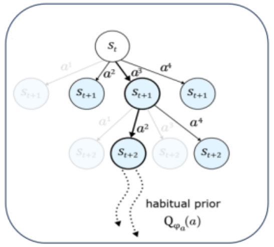
Algorithm 1 DAIMC on-policy training

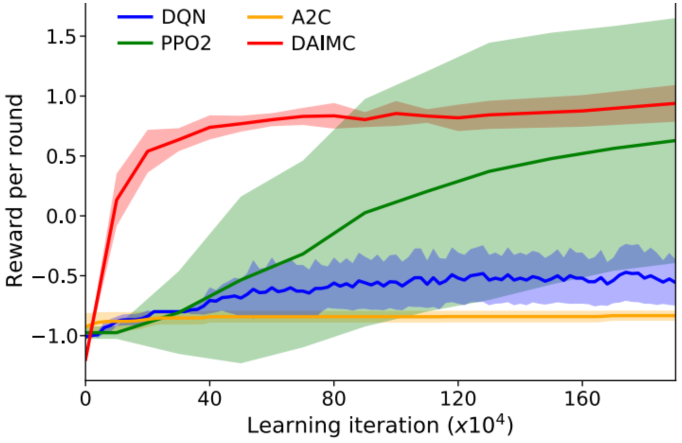
for $t = 1, 2, \ldots,$ max iterations **do**
  Randomize environment and sample a new observation $\tilde{o}_t$.
  Run planner and compute prior policy $P(a_t)$.
  Compute $Q_{\phi_s}(s_t)$ using $\tilde{o}_t$.
  Compute $Q_{\phi_a}(a_t)$ using a sampled state $\tilde{s}_t \sim Q_{\phi_s}(s_t)$.
  Compute $D_t = D_{\mathrm{KL}}\left[Q_{\phi_a}(a_t) \parallel P(a_t)\right]$.
  Apply a gradient step on ϕ_a using D_t as loss.
  Compute ω_{t+1} from Eq. (10) using D_t.
  Apply action $\tilde{a}_t \sim P(a_t)$ to the environment and sample a new observation $\tilde{o}_{t+1}$.
  Compute μ, σ from $P_{\theta_s}(s_{t+1}|\tilde{s}_t, \tilde{a}_t)$.
  Compute $Q_{\phi_s}(s_{t+1})$ using $\tilde{o}_{t+1}$.
  Apply a gradient step on θ_s using $D_{\mathrm{KL}}\left[Q_{\phi_s}(s_{t+1}) \parallel \mathcal{N}(\mu, \sigma^2/\omega_t)\right]$.
  Apply a gradient step on ϕ_s, θ_o using $-\mathbb{E}_{Q(s_{t+1})}\left[\log P_{\theta_o}(o_{t+1}|s_{t+1})\right] + D_{\mathrm{KL}}\left[Q_{\phi_s}(s_{t+1}) \parallel \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2/\omega_t)\right]$.
end for



Planning: Monte-Carlo Tree Search + Habit



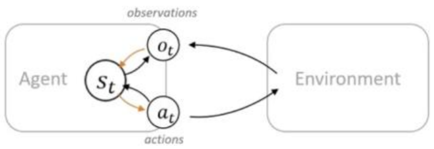


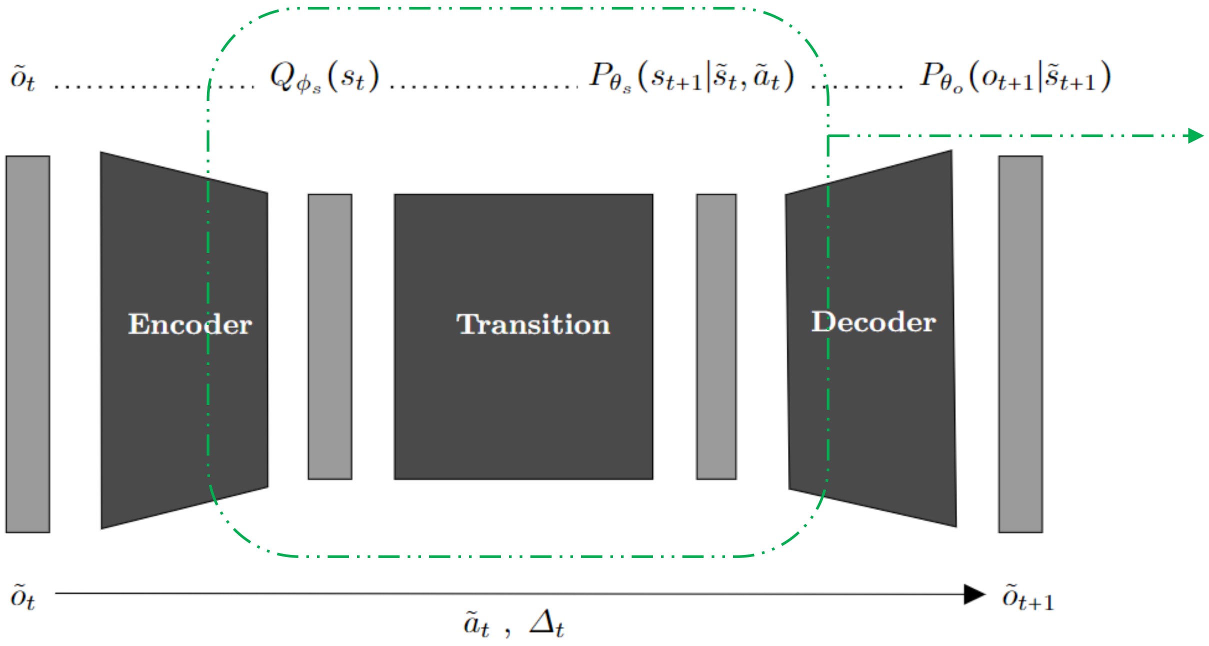


[4] Fountas et al., 2020

Active Inference: Core

The core of the formalism is the Generative Model.
Realization of this formalism is through Deep Learning: Generative Modeling and Inferences





Variational Auto Encoder

World-Modelling and Model-Based RL

Model-Based RL

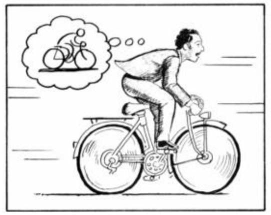
- Use a model for making decisions
- Engaging methods for policy improvement and planning

World Models

- Build an internal predictive/generative model of the world (often task agnostic)
- Focuses on modelling and representations, either latent or ambient
- Learning policy (or planning) in model imaginations (roll outs)

General agents contain world models

Jonathan Richens[1] David Abel[1] Alexis Bellot[1] Tom Everitt[1]



Abstract

Are world models a necessary ingredient for flexible, goal-directed behaviour, or is model-free learning sufficient? We provide a formal answer to this question, showing that any agent capable of generalizing to multi-step goal-directed tasks must have learned a predictive model of its environment. We show that this model can be extracted from the agent's policy, and that increasing the agents performance or the complexity of the goals it can achieve requires learning increasingly accurate world models. This has a number of consequences: from developing safe and general agents, to bounding agent capabilities in complex environments, and providing new algorithms for learning an agent's world model.

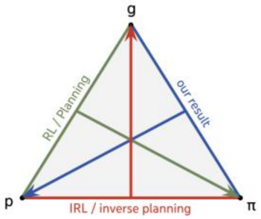
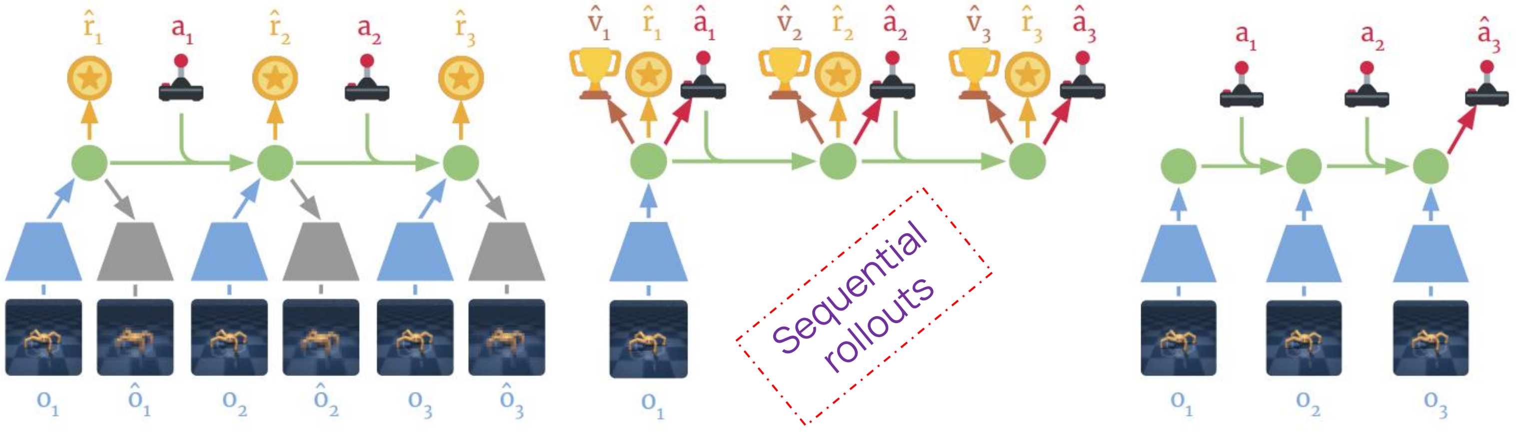


Figure 1: Our result complements previous insights from planning and inverse RL. While planning uses a world model and a goal to determine a policy, and IRL and inverse planning use an agent's policy and a world model to identify its goal, our result uses an agent's policy and its goal to identify a world model



(a) Learn dynamics from experience (b) Learn behavior in imagination (c) Act in the environment

RSSM

Sequence model:	$h_t = f_\phi(h_{t-1}, z_{t-1}, a_{t-1})$
Encoder:	$z_t \sim q_\phi(z_t \mid h_t, x_t)$
Dynamics predictor:	$\hat{z}_t \sim p_\phi(\hat{z}_t \mid h_t)$
Reward predictor:	$\hat{r}_t \sim p_\phi(\hat{r}_t \mid h_t, z_t)$
Continue predictor:	$\hat{c}_t \sim p_\phi(\hat{c}_t \mid h_t, z_t)$
Decoder:	$\hat{x}_t \sim p_\phi(\hat{x}_t \mid h_t, z_t)$

Dreamer learns policy with Actor-Critic in model imaginations [5]

[5] Hafner et al., 2023.

Motivation

Machine Learning

Concept/Formalism	Architeture/Algorithm	KPI Improvement
Theoretical Development	**Model Development**	**Application Development**

Promises of Deep Learning: Generative Models:

- *Bayesian and Probabilistic*
- *Representation Learning for both latent and ambient spaces*
- *Utilizes the observatory (cheap and raw) data*

Promises of Active Inference (AIF):

- *Bayesian Framework: Integrates perception, learning, and action*
- *Model-Driven*
- *Adaptability*
- *Intrinsic Objectives: Exploration to resolve uncertainty*
- *End-to-End Action: Doesn't require high-quality reward signals*

Promises of Word-Model RL Agents:

- *Planning via imagination*
- *Generalization & Sample-efficient*
- *Reduced reliance on rewards*

Problem:

AIF agents struggle in delayed and long-horizon environments due to reliance on:
- *Immediate predictions.*
- *Exhaustive planning.*

Train a Policy within the Generative World-Model via Active Inference
Conencting: DL, AIF, RL

Overshooting Trick: Integrating Policy

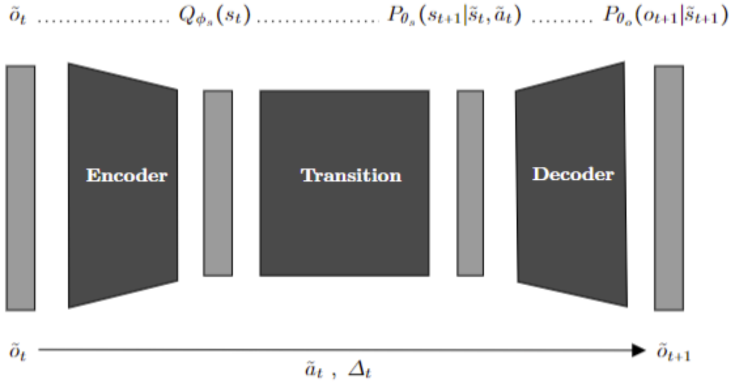


Figure 3: The agent's architecture and generative framework resemble that of a VAE. The line on top represents the agent simulating the future and making a prediction, while on the bottom, the agent receives a new observation after Δ_t of taking an action, $\tilde{a}_t$.

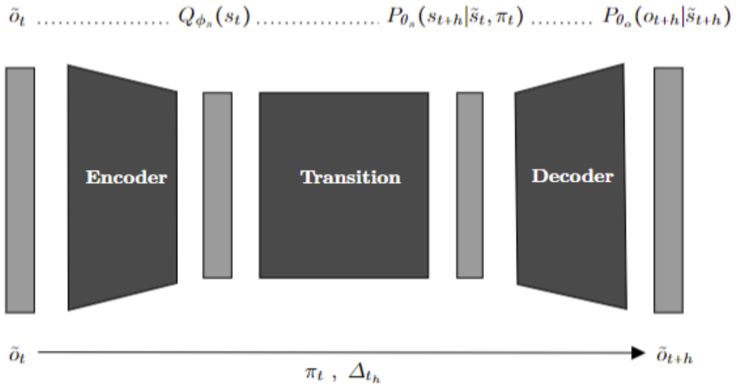
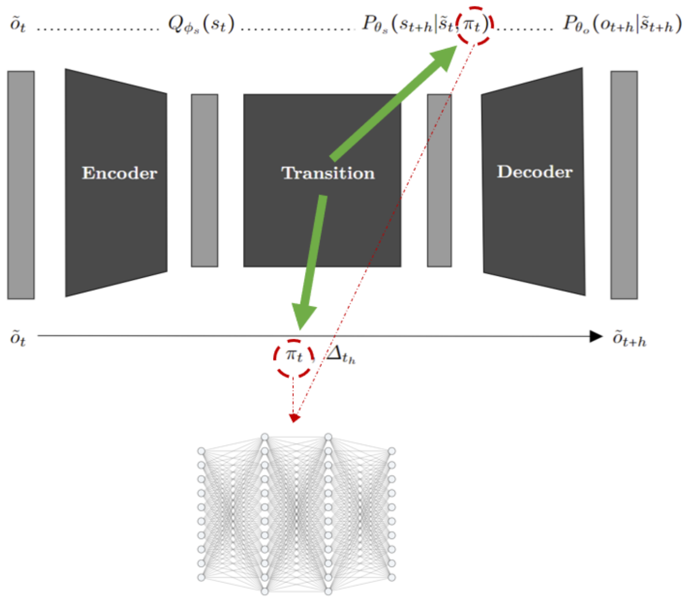


Figure 4: The agent's architecture and generative framework with an an impact/update/planning horizon, h (e.g., 100).



Selecting an impact/update/planning horizon: h (e.g., 300)

Formalism: Deep Active Inference (DAIF)

Model Learning (Calibration): Variational Free Energy

$$\theta^* = \arg\min_{\theta} \left(\mathbb{E}_{Q_\phi(s_t, a_t)} \left[\log Q_\phi(s_t, a_t) - \log P_\theta(o_t, s_t, a_t) \right] \right)$$

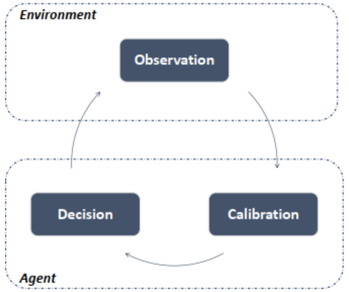
Decision (Planning & Action): Expected Free Energy

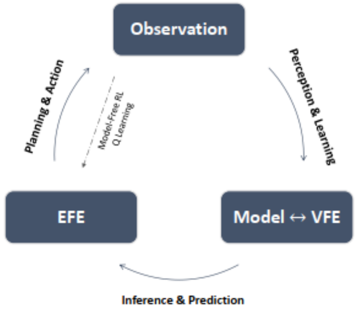
$$\begin{aligned} G(\pi,\tau) = & \boxed{-\mathbb{E}_{Q(\theta|\pi)Q(s_\tau|\theta,\pi)Q(o_\tau|s_\tau,\theta,\pi)} \left[\log P(o_\tau|\pi) \right]} \quad 1 \\ & \boxed{+\mathbb{E}_{Q(\theta|\pi)} \left[\mathbb{E}_{Q(o_\tau|\theta,\pi)} H(s_\tau|o_\tau,\pi) - H(s_\tau|\pi) \right]} \quad 2 \\ & \boxed{+\mathbb{E}_{Q(\theta|\pi)Q(s_\tau|\theta,\pi)} H(o_\tau|s_\tau,\theta,\pi) - \mathbb{E}_{Q(s_\tau|\pi)} H(o_\tau|s_\tau,\pi)} \quad 3 \end{aligned}$$

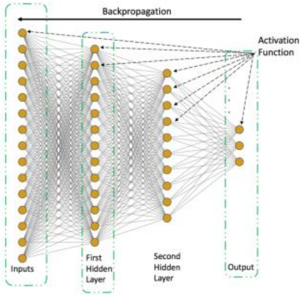
1) Extrinsic Value 2) State Epistemic Uncertainty 3) Parameter Epistemic Uncertainty

Policy Optimization via Gradients

1. Integrate the encoded policy parameters $\hat{\pi}(\phi_a)$ in the model and EFE: $G_\theta(\tilde{o}, \phi_a)$
2. Take the gradient of EFE w.r.t parameters: $\phi_a \leftarrow \phi_a - \alpha \nabla_{\phi_a} \mathbb{E}\left[G(\phi_a) \right]$



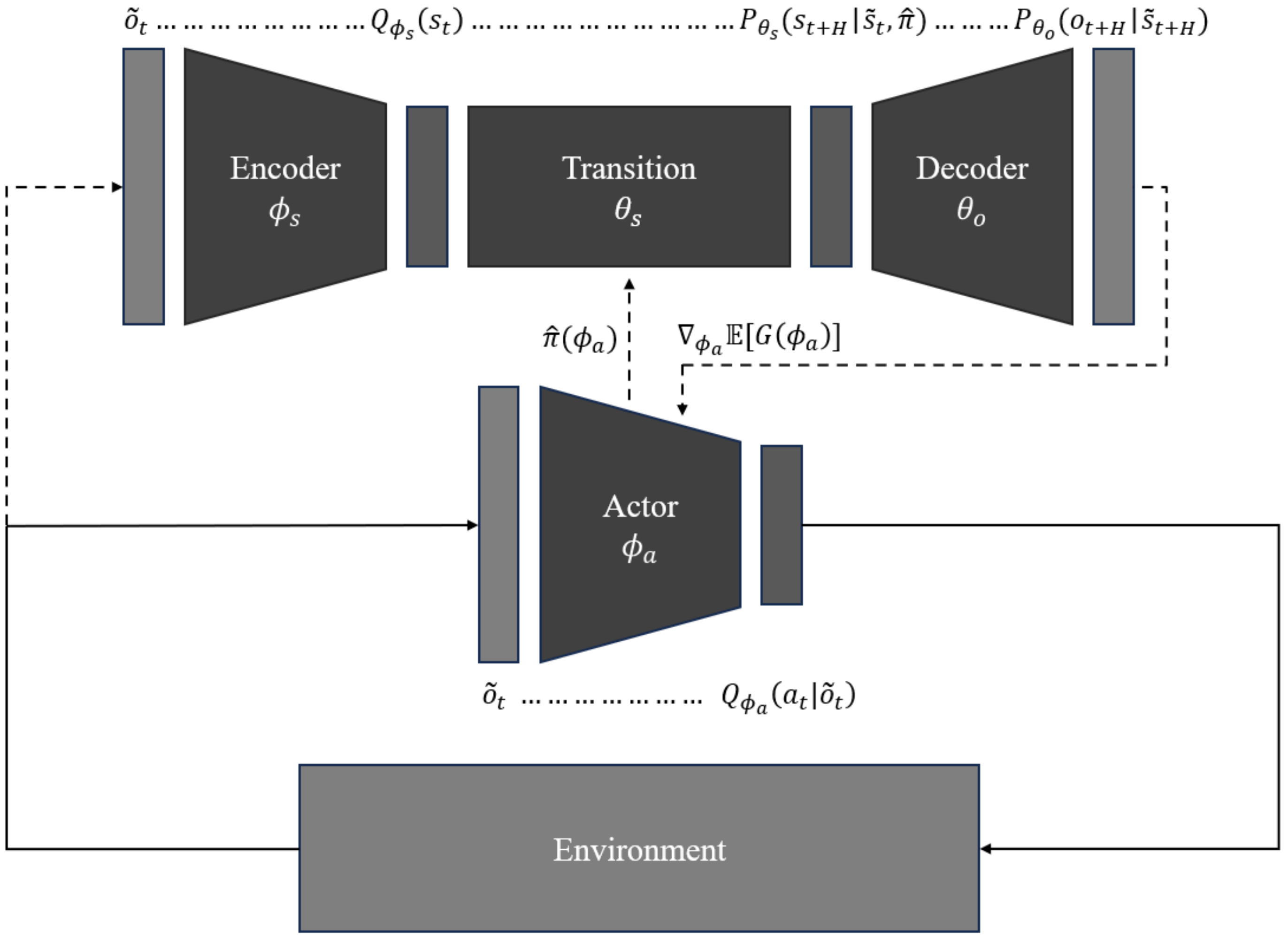




$$\Pi : \mathcal{Q}_{\phi_a} \rightarrow \hat{\pi}.$$

DAIF: Architecture

- Actor directly interact with the environment.
- Actor parameters are encoded and integrated into transition.
- Model predicts H steps into future in a single rollout.
- Gradient of policy parameters are taken via EFE of the prediction of the H steps into future.



DAIF: Algorithm

- Actor directly interact with the environment.
- Actor parameters are encoded and integrated into transition.
- Model predicts H steps into future in a single rollout.
- Gradient of policy parameters are taken via EFE of the prediction of the H steps into future.

Algorithm 1 Deep AIF Agent Training (per epoch)

Initialize $\theta = \{\theta_s, \theta_o\}$, $\phi = \{\phi_s, \phi_a\}$, $\mathcal{M}$

Randomly initialize E

for $n = 1, 2, ..., N$ **do**

▷ ENVIRONMENT INTERACTION

$\hat{\pi}_t \leftarrow \Pi(Q_{\phi_a})$

for $\tau = t+1, t+2, ..., t+H$ **do**

Sample a new observation $\tilde{o}_\tau$ from E

Apply $\tilde{a}_\tau \sim Q_{\phi_a}(a_\tau | \tilde{o}_\tau)$ to E

Sample a new observation $\tilde{o}_{\tau+1}$ from E

$\mathcal{M} \leftarrow \mathcal{M} \cup \{(\tilde{o}_t, \hat{\pi}_t, \tilde{o}_{t+H})\}$

▷ MODEL LEARNING

$\{(\tilde{o}_{t'}, \hat{\pi}_{t'}, \tilde{o}_{t'+H})\}^{B_1} \sim \mathcal{M}$

for $t' = 1, 2, ..., B_1$ **do**

run Model$(\tilde{o}_{t'}, \hat{\pi}_{t'}, \tilde{o}_{t'+H})$

$\mathcal{L}_s \leftarrow \mathcal{L}_s + D_{\mathrm{KL}}\left[Q_{\phi_s}(s_{t'+H}) \,||\, \mathcal{N}(\mu, \sigma^2)\right]$

$\mathcal{L}_o \leftarrow \mathcal{L}_o - \mathbb{E}_{Q(s_{t'+H})}\left[\log P_{\theta_o}(o_{t'+H} | \tilde{s}_{t'+H})\right]$

$\mathcal{L}_o \leftarrow \mathcal{L}_o + \beta * D_{\mathrm{KL}}\left[Q_{\phi_s}(s_{t'+H}) \,||\, \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)\right]$

$\theta_s \leftarrow \theta_s - \xi \nabla_{\theta_s} \mathbb{E}\left[\mathcal{L}_s(\theta_s)\right]$

$\phi_s \leftarrow \phi_s - \gamma \nabla_{\phi_s} \mathbb{E}\left[\mathcal{L}_s(\phi_o)\right]$

$\theta_o \leftarrow \theta_o - \eta \nabla_{\theta_o} \mathbb{E}\left[\mathcal{L}_o(\theta_o)\right]$

▷ POLICY OPTIMIZATION

$\{\tilde{o}_\tau\}^{B_2} \sim \mathcal{M}$

for $\tau = 1, 2, ..., B_2$ **do**

Compute $Q_{\phi_s}(s_\tau)$ using $\tilde{o}_\tau$

Sample $\tilde{s}_\tau \sim Q_{\phi_s}(s_\tau)$

for $s = 1, 2, ..., S_1$ **do**

Compute $\mu, \sigma \leftarrow P_{\theta_s}(s_{\tau+H} | \tilde{s}_\tau, \hat{\pi}_t)$

Sample $\tilde{s}_{\tau+H} \sim \mathcal{N}(\mu, \sigma^2)$

Compute $P_{\theta_o}(o_{\tau+H} | \tilde{s}_{\tau+H})$

Compute $Q_{\phi_s}(\tilde{s}_{\tau+H})$ using $\tilde{o}_{\tau+H}$

Compute $\mu', \sigma' \leftarrow Q_{\phi_s}(\tilde{s}_{\tau+H})$

$G \leftarrow G - \log \Psi\left[P_{\theta_o}(o_{\tau+H} | \tilde{s}_{\tau+H})\right]$

$G \leftarrow G + \left[H(\mu', \sigma') - H(\mu, \sigma)\right]$

for $s = 1, 2, ..., S_2$ **do**

Sample $\tilde{s}_{\tau+H} \sim P_{\theta_s}(s_{\tau+H} | \tilde{s}_\tau, \hat{\pi}_\tau)$ ▷ *Re-computed with dropout.*

Compute $\mu'', \sigma'' \leftarrow P_{\theta_o}(o_{\tau+H} | \tilde{s}_{\tau+H})$

Sample $\tilde{s}_{\tau+H} \sim \mathcal{N}(\mu, \sigma^2)$

Compute $\mu''', \sigma''' \leftarrow P_{\theta_o}(o_{\tau+H} | \tilde{s}_{\tau+H})$

$G \leftarrow G + \left[H(\mu'', \sigma'') - H(\mu''', \sigma''')\right]$

$\phi_a \leftarrow \phi_a - \alpha \nabla_{\phi_a} \mathbb{E}\left[G(\phi_a)\right]$

Agent components:

Model:

Encoder Q_{ϕ_s}.

Transition P_{θ_s}.

Decoder P_{θ_o}.

Actor Q_{ϕ_a}.

Actor mapping Π.

Preference mapping Ψ.

Other components:

Environment E.

Experience Memory $\mathcal{M}$.

Hyperparameters:

Iterations N.

Beta β.

Horizon H.

Batch size B_1 , B_2.

Sample size S_1 , S_2.

Learning rate ξ , γ , η , α.

Run Model$(\tilde{o}_i, \hat{\pi}, \tilde{o}_{i+H})$:

Compute $Q_{\phi_s}(s_i)$ using $\tilde{o}_i$

Sample $\tilde{s}_i \sim Q_{\phi_s}(s_i)$

Compute $\mu, \sigma \leftarrow P_{\theta_s}(s_{i+H} | \tilde{s}_i, \hat{\pi})$

Compute $Q_{\phi_s}(\tilde{s}_{i+H})$ using $\tilde{o}_{i+H}$

Compute $\mu', \sigma' \leftarrow Q_{\phi_s}(\tilde{s}_{i+H})$

Sample $\tilde{s}_{i+H} \sim \mathcal{N}(\mu, \sigma^2)$

Compute $P_{\theta_o}(o_{i+H} | \tilde{s}_{i+H})$

Benchmark: AIF vs Realistic Industrial Environment

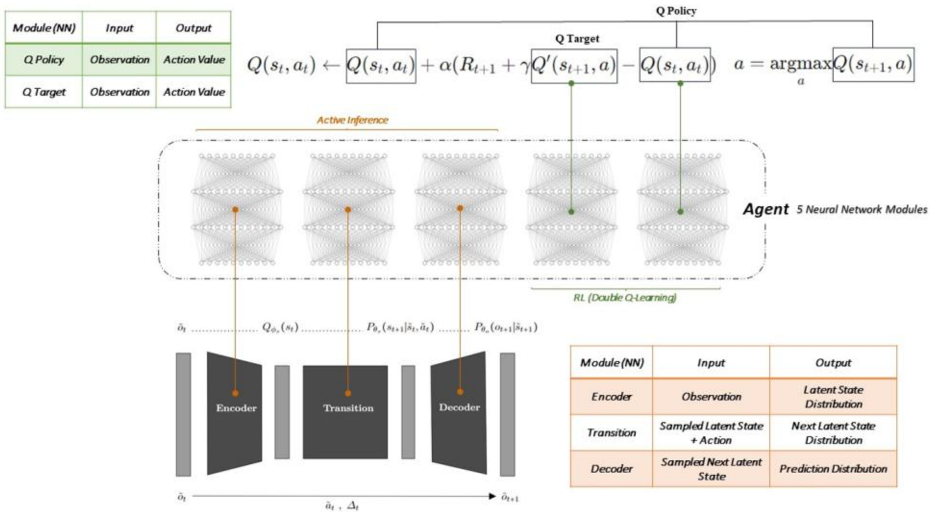
Application

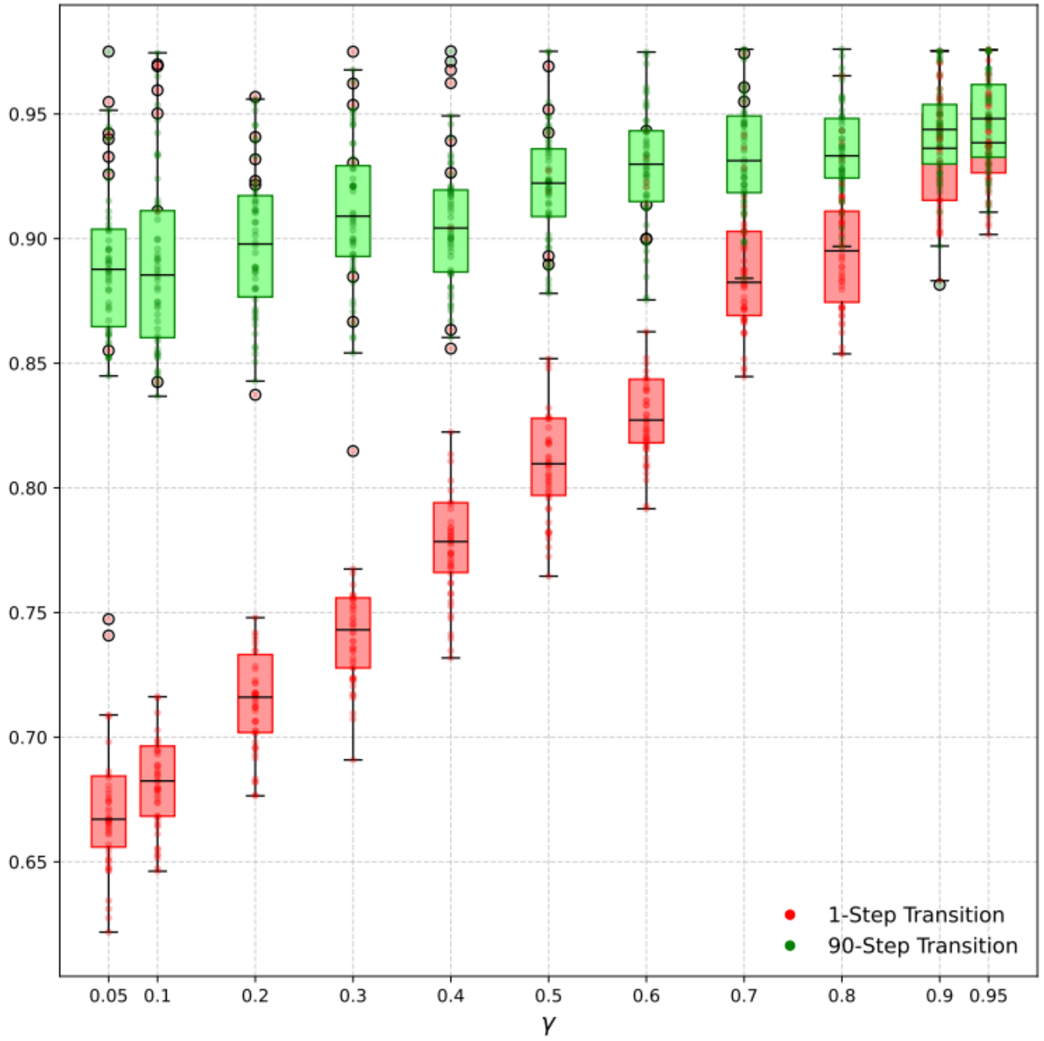
Energy-Efficient Control of simulated workstations within automotive manufacturing system composed of parallel, identical machines.

Challenges

Stochastic - Delayed - Long-horizon - Multi-Modal Observation

Requires extensive planning with horizon of one-shift (~ 3000 actions).





Active Inference

Reinforcement Learning

Results: DAIF vs Realistic Industrial Environment

Application

Energy-Efficient Control of simulated workstations within automotive manufacturing system composed of parallel, identical machines.

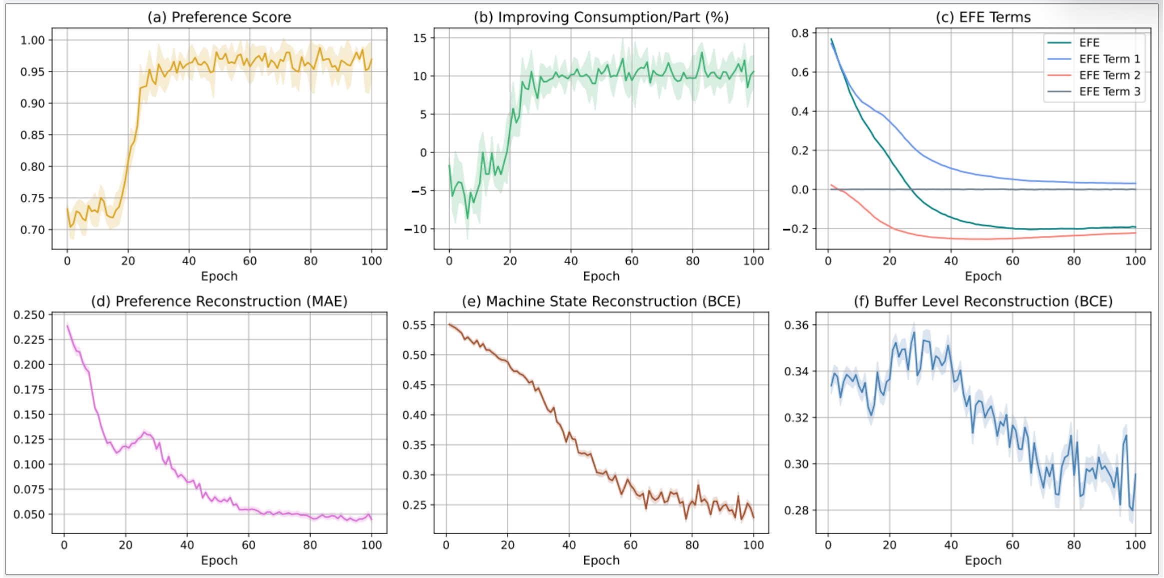
Challenges

Stochastic - Delayed - Long-horizon - Multi-Modal Observation

Requires extensive planning with horizon of one-shift (~ 3000 actions).

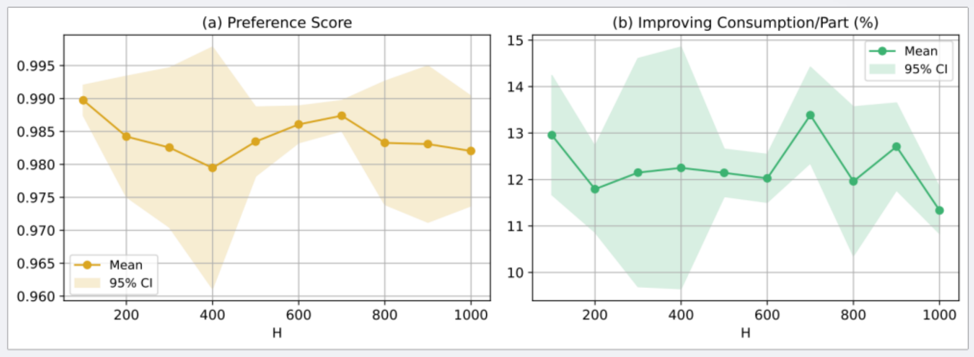
Improving Energy-Efficiency:

Controlling the number of active machines to minimize idle time with negligible production loss.



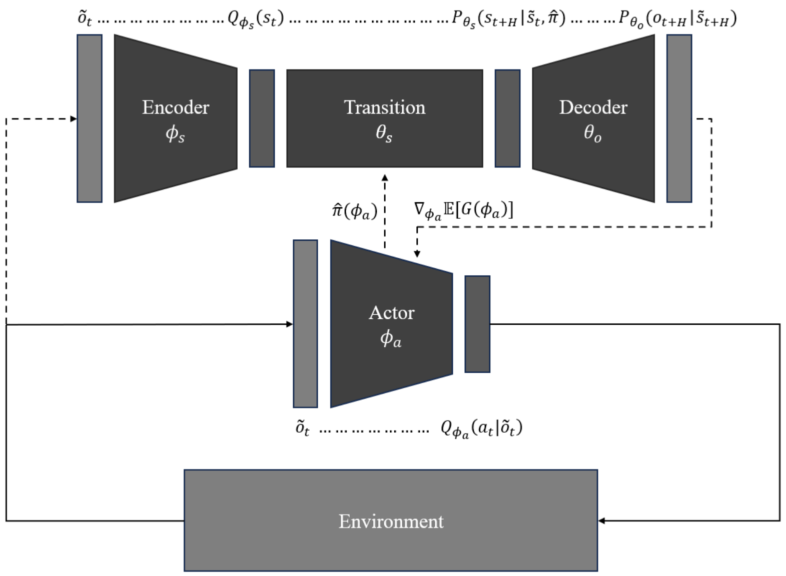
Agent(ϕ)	Production Loss [%]	EN Saving [%]
DQN (0.93)	4.82 ± 0.34	10.87 ± 0.76
DQN (0.94)	3.34 ± 0.23	9.92 ± 0.69
DAIF	$\mathbf{2.59 \pm 0.16}$	$\mathbf{12.49 \pm 0.04}$
DQN (0.95)	1.27 ± 0.05	7.00 ± 0.07
DQN (0.96)	1.27 ± 0.09	7.62 ± 0.12
DQN (0.97)	1.20 ± 0.05	7.72 ± 0.10
DQN (0.98)	0.54 ± 0.04	2.72 ± 0.19
DQN (0.99)	0.40 ± 0.03	2.46 ± 0.01

Production loss versus energy-saving (EN) across reward parameters ϕ of DQN agents (best model-free RL) against the DAIF agent .



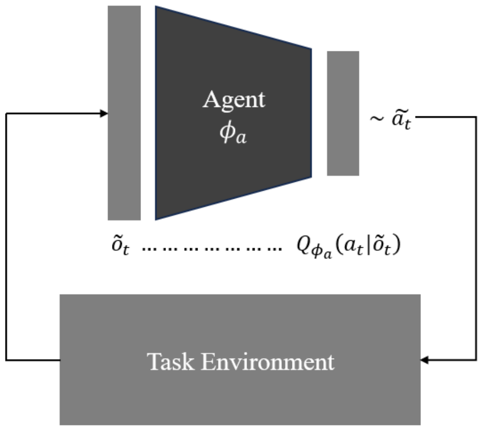
Performance of the agents versus overshooting horizon H.

DAIF: Promises



- Derived from a Bayesian-grounded framework (heuristic-free)
- Natively scalable to both discrete and continuous action spaces
- A unified explore–exploit gradient
- Effective encoding of long-horizon dynamics
- Capturing distinct forms of epistemic uncertainty, providing an intrinsic drive for model refinement
- Reliance on observations (cheap, raw data), instead of often expensive, engineered reward signals
- A degree of interpretability
- Inherent adaptability in non-stationary settings via AIF
- Minimal computational cost during both inference (model-free) and planning (EFE calculated in a single H-step forward pass)

DAIF: Potential



Scope: Data-Driven Decision Making or Optimization:

Sequential decision making under uncertainty (POMDP)

- Control / Planning / Search
- Reasoning / Inference / Search
- Generation / Synthesis

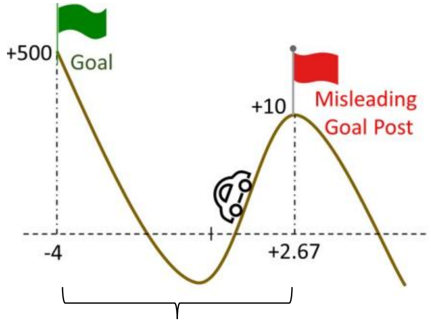
Domains:

- Robotics
- Production Control
- Healthcare
- Drug Discovery / Protein Design
- …

Settings:

Delayed and Long-Horizon Environments

- Stochastic & Partial Observability
- Delayed/Sparse/Expensive Reward (i.e., Scarce high-quality data)
- Long Horizon



Two sparse rewards [1]

[1] Chakraborty et al., 2022

Future Work

DAIF is a concept: Generative world model with multi-step latent transitions + differentiable policy; backpropagates EFE through long horizons without tree search; scales to all spaces while keeping the explore–exploit balance.

Next steps

- Generative Model (DL): Try diffusion/flow-matching world models instead of VAE.
- Sample efficiency (DL): Still needs experience collection every H steps; aim to reduce data needs.
  - Sequence aggregation: Replace slow recurrent unrolling with set-based pooling of H embeddings (with simple positional encodings) or operator-learning for resolution-invariant aggregation.
- Optimization (RL): Explore actor–critic training and add regularization to stabilize/variance-reduce EFE gradients.
- EFE Methods (AIF): Improving distinct term estimations (e.g., sampling).
- Adaptation (AIF): Focus on rapid learning in non-stationary environments.
- Expanding Experimental: In addition to more domains, looking a range of settings.
- Reasoning Model: Exploring potential for reasoning tasks.
- Big picture: Bridges generative world-modelling with active inference and RL—compact, end-to-end probabilistic agents

References

Yeganeh, Yavar Taheri, Mohsen Jafari, and Andrea Matta. "Deep Active Inference Agents for Delayed and Long-Horizon Environments." *arXiv preprint arXiv:2505.19867* (2025).

Paper

Code

[1] Chakraborty, Souradip, et al. "Dealing with sparse rewards in continuous control robotics via heavy-tailed policies." arXiv preprint arXiv:2206.05652 (2022).
[2] Lanillos, P., Meo, C., Pezzato, C., Meera, A. A., Baioumy, M., Ohata, W., ... & Tani, J. (2021). Active inference in robotics and artificial agents: Survey and challenges. arXiv preprint arXiv:2112.01871.
[3] Huang, Y., Li, Y., Matta, A., & Jafari, M. (2024). Navigating Autonomous Vehicle on Unmarked Roads with Diffusion-Based Motion Prediction and Active Inference. arXiv preprint arXiv:2406.00211.
[4] Fountas, Z., Sajid, N., Mediano, P., & Friston, K. (2020). Deep active inference agents using Monte-Carlo methods. Advances in neural information processing systems, 33, 11662-11675.
[5] Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2023). Mastering diverse domains through world models. arXiv preprint arXiv:2301.04104.