Learning to Play with Intrinsically-Motivated Self-Aware Agents

Nick Haber1,2,3*, Damian Mrowca4*, Stephanie Wang4, Li Fei-Fei4, Daniel L. K. Yamins1,4,5

Department of Psychology1, Pediatrics2, Biomedical Data Science4, Computer Science4, and Wu Tsai Neurosciences Institute6, Stanford, CA 94305

Introduction

Emergent Behaviors

Planning

Task Transfers

Future Work

A sequence of behaviors emerges:

- Latent Future prediction
- Random Policy
- Random World model
- Self-aware Policy
- Inverse Dynamics prediction

Emerging behaviors across different models. Different combinations of world-model tasks and policy mechanisms are compared.

This work was supported by grants from the James S. McDonnell Foundation, Simons Foundation, and Sloan Foundation (DLKY), a Berry Foundation postdoctoral fellowship and Stanford Department of Biomedical Data Science (NLH-15-LT17010533-31, NLH-14-LTR0190051), and NSF (Grant 1450164 - 16-12127 and CNN - P1R3 - SRIR (UCLA lead) -155 (GT275-57)).

Approach: Self-Model Vs World-Model

Task Transfers

Model comparison on task transfers including object presence (binary classification), localization (pixel-wise 3D positional vector), and categorization (16-way categorization). Linear estimators were trained on top of the output features of each model.

This was a more photorealistic simulation environment. Next steps include:

- more photorealistic simulation
- more realistic agent embodiment
- curiosity towards the novel but learnable
- animate attention and theory of mind
- comparision to human developmental data

Curiosity-based learning mechanism

- The world model (blue) solves a dynamics prediction problem.
- The self-model (red) seeks to predict the world-model’s loss and is learned simultaneously.
- Actions are chosen to antagonize the world-model, leading to novel and surprising events in the environment (black).
- This creates a virtuous cycle in which the agent chooses novel but predictable actions.
- Playful behavior emerges as the agent pushes the boundaries of what its world-model-prediction systems can achieve.
- As world-modeling capacity improves, what used to be novel becomes old hat, and the cycle repeats.

Approach: Self-Model Vs World-Model

Emergent Behaviors

Planning

Future Work

Intrinsically-motivated self-aware architecture.