Embedding vision models in RL environments

To study how ecology constrains visual processing, we present a reinforcement learning (RL) framework in which an agent aims to survive in a 3-d environment that it perceives through a vision model. Our code is built on:

- PyTorch, the DNN library we use to implement vision models,
- Sample Factory [2], a deep RL library that we use to train agents, and
- ViZDoom [3], an RL environment engine based on Doom.

Varying the visual complexity of a task

The agent has “satiety” from 0 to 100, and continuously loses satiety. To survive and thrive, it must solve a visually-guided foraging task.

- Nourishment and poison are randomly distributed in the environment.
- We modulate task difficulty by varying the diversity of the nourishment/poison.

RL Primer: We describe the world in terms of states $s$, actions $a$, and rewards $r$.

- transition distribution: $p(s′ | s, a)$
- policy: $π(s) = p(a | s)$
- action-value function: $Q(s, a) = E[\sum_{t=0}^{\infty} γ^t R_t | S_0 = s, A_0 = a]$;
- policy gradient theorem: $\frac{1}{N} \sum_{n=1}^{N} \log π(a_n | s_n) \frac{∂Q(s, a)}{∂θ}$
- Linear regression on the estimated value function $\hat{V}$ shows the importance of environmental variables to its representation of the world.

- $\hat{V}(s) = E[\sum_{t=0}^{∞} γ^t R_t | S_0 = s]$.

Brain complexity scales with visual complexity

Satiety signals enable unique behaviours

- Satiety signals allow agents to strategize around nourishment magnitude.
- Wasted nourishment occurs when nourishments yields $100 > 0$.

Brain architecture drives different representations

- Linear regression on the estimated value function $\hat{V}$ shows the importance of environmental variables to its representation of the world.
- $\hat{V}(s) = E[\sum_{t=0}^{∞} γ^t R_t | S_0 = s]$.

Conclusion and Outlook

We have laid foundations for modelling ecological constraints on vision and shown:

- how to bridge established results from DNN vision models [1] into a dynamic RL setting,
- that rich behaviours and representations emerge even on simple tasks, and
- that brain architecture and visual complexity interact in nontrivial ways.

Building on this, our next goals are to explore:

- more complex environmental objects such as predators,
- adding metabolic costs for the magnitude of network activity to induce “implicit” efficient coding, and
- modelling tradeoffs between minimizing image reconstruction error and task-relevant efficient codes.

References


A computational approach to visual ecology with deep reinforcement learning

Sacha Sokoloski1,2 and Philipp Berens1,2

1. Hertie Institute for AI in Brain Health, Tübingen; 2. University of Tübingen