Emergent active sensing behaviors in artificial electric fish agents

Sonja Johnson-Yu[1], Satpreet H. Singh[1], Federico Pedraja[2], Denis Turcu[2], Pratyusha Sharma[3], Naomi Saphra[1], Nathaniel B. Sawtell[2], Kanaka Rajan[1] (kanaka_rajan@hms.harvard.edu)
Harvard University[1], Columbia University[2], MIT[3]

Abstract

Weakly electric fish, such as *Gnathonemus petersii*, generate pulsatile electric organ discharges (EODs) that enable them to sense their environment through active electrolocation. This plays a crucial role in several key behaviors, such as navigation, foraging, and avoiding predators. While the anatomical and physiological organization of the active electrosensory system has been extensively studied, the contribution of active electrolocation to adaptive behavior in naturalistic settings remains relatively underexplored. Here we present a preliminary *in silico* model of active sensing in electric fish, using a neural network-based artificial agent trained by deep reinforcement learning to perform an analogous active sensing task in a 2D environment. The trained agent recapitulates key features of natural EOD statistics, shows emergent behavioral modularity, and provides intuitions about the representation of key latent variables governing agent behavior, such as energy levels (satiety).

Keywords: active sensing; weakly electric fish; deep reinforcement learning; artificial agents; recurrent neural networks;

Introduction

Weakly electric fishes like *Gnathonemus petersii* use electric pulses, or electric organ discharges (EODs), to actively sense their environment, communicate with each other, and sense their environment based on the EODs of nearby fish (Von der Emde, 1999; Sawtell, Williams, & Bell, 2005; Pedraja & Sawtell, 2024). The role that active electrolocation plays in the goal-oriented behaviors of fish is less well understood compared to our extensive knowledge of the physiology of the neural mechanisms responsible for EOD generation. This knowledge gap is due to the difficulty of designing naturalistic yet well-controlled studies that capture the complexity of the animals' sensory ecology and behavioral repertoire.

In recent years, neural network-based artificial agents trained to perform different tasks have emerged as powerful tools to model animal behaviors and neural computations (Haesemeyer, Schier, & Engert, 2019; Singh, van Breugel, Rao, & Brunton, 2023). By transforming sensory inputs into motor outputs similar to those of real animals, such models offer insight into the neural and cognitive processes underlying animal behaviors. They also enable flexible *in silico* experimentation while being fully observable, allowing hypothesis testing where experimental data collection is challenging.

Here, we present preliminary results from a biologically-constrained artificial agent trained by deep reinforcement learning (DRL) to perform an active-sensing foraging task in a 2D environment, analogous to weakly electric fish behavior.

Environment and Agent

Inspired by lab experiments on *Gnathonemus petersii*, we train our agents in simulated 2D tanks of size 60 cm x 60 cm (Fig 1a). Simulations are initialized with n food items placed uniformly at random. The position and orientation of a single agent are also initialized uniformly at random. At each timestep, the agent observes the egocentric vector distance of the nearest food item and the nearest wall within its sensing range. It can also observe its internal energy levels ($e \in [0,1]$), which increase every time it eats food, but otherwise decrease linearly with time and activity levels. At each timestep, it decides how much to move forward, how much to turn, and whether or not to emit an EOD. The agent is rewarded for eating food, penalized for both starvation and overeating, and has a baseline metabolic cost associated with staying alive. The agent (Fig 1a, right) consists of a recurrent neural network (RNN) (Rajan, Harvey, & Tank, 2016) followed by parallel two-layer Actor and Critic Multi-Layer Perceptrons (MLPs). The former selects the agent's actions, and the latter estimates the value of actions during training using policy gradients (Ni, Eysenbach, & Salakhutdinov, 2021). All layers are 64-units wide, with *tanh* nonlinearities. We constrain the agent's maximum linear and angular velocities and accelerations to match experimental data collected from an electric fish in an identically-sized tank. Simulations are run at $\approx$ 83 FPS to enable a minimum SPI of 12ms, as is observed in lab experiments. For simplicity, here, the agent actions and observations are deterministic.

Results

Trained agents successfully electrolocate food items while producing movement trajectories (Fig. 1b) and EOD transcripts (Fig. 1c) that resemble experimental data from real fish. Two behavior modes (Fig 1d and e), namely "resting" and "active foraging", are also observed, similar to those observed in real fish (von der Emde, 1992). We also observe that the trained agent learns a "homeostatic drive" to maintain its energy levels slightly below the maximum possible (Fig. 1e). Additionally, we find that the high energy (satiated) state is correlated with low linear velocity and vice versa (Fig. 1e). Furthermore, the low- and high- energy modes are observable in the RNN's hidden state activities (Fig. 1f).

Future Work

Next, we plan to explore the role of EODs in more complex tasks involving cooperation and competition between multiple identical and diverse agents (conspecifics and heterospecifics), along with richer biophysical models incorporating mechanisms of EOD generation and reception.

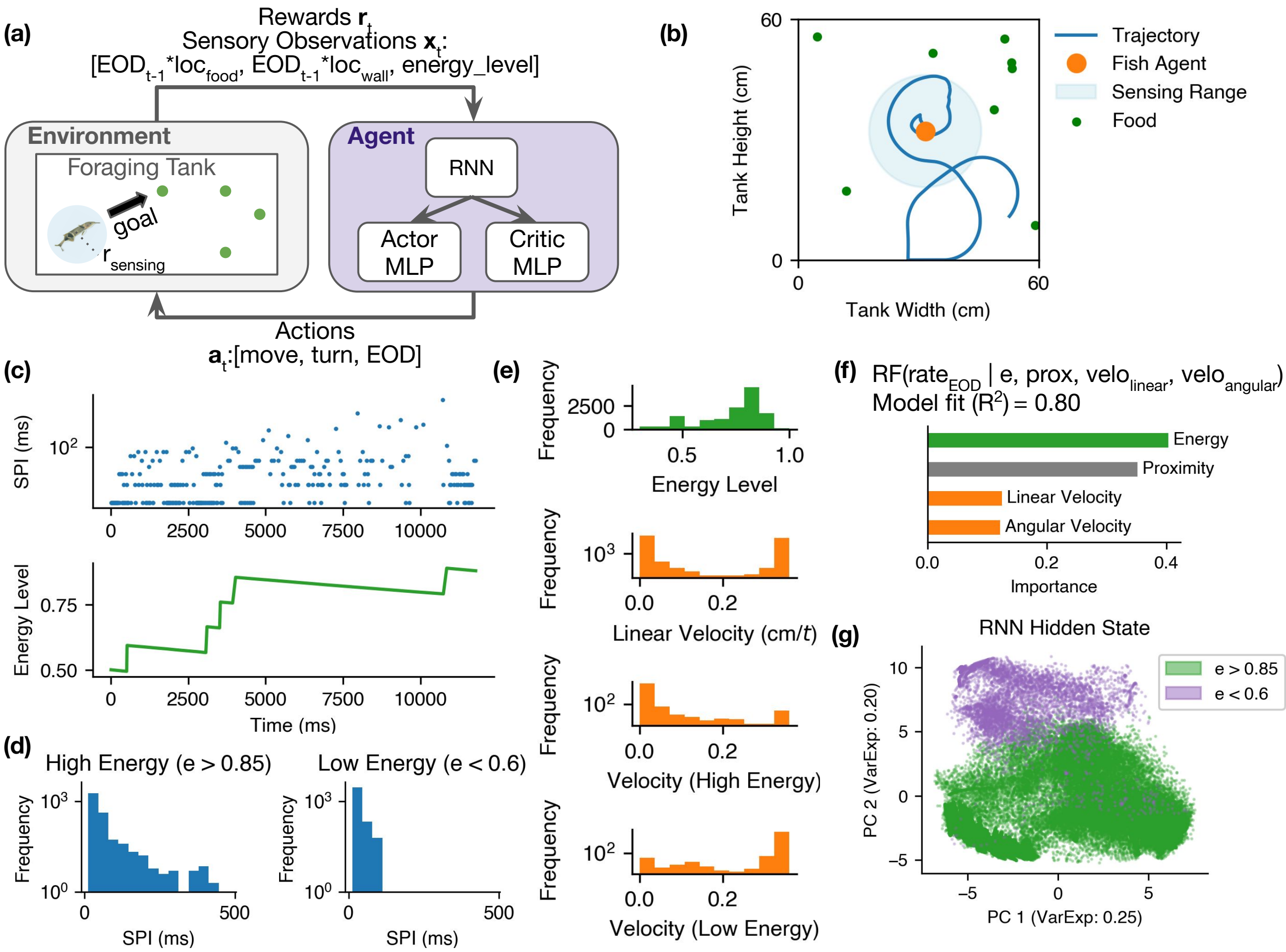

Figure 1: **(a) The agent works in tandem with the tank simulation environment to learn an efficient foraging policy** (billycorgan84, 2009). At each timestep, the agent receives sensory observations and rewards from the environment and then selects its actions. If the agent emitted an EOD in the previous timestep, it can observe the location of the nearest food and wall within its sensing range. The agent uses a recurrent neural network (RNN) to infer the environmental ('belief') state, selects an action using an Actor multilayer perceptron (MLP), and estimates the action's value with a Critic MLP. **(b) Example trajectory from a trained agent**. Food is distributed uniformly at random throughout the 60cm $\times$ 60cm tank. The agent can sense food within its sensing range (radius=14cm). **(c) Example sequential pulse intervals (SPIs) and energy of a trained agent over a 1500-timestep episode**. A sequential pulse interval is the length of time between two EODs. Periods of repeated low SPIs (frequent EODs) correlate with vigorous foraging behavior, seen in the step-wise increases in energy between 0-5000ms. High energy (satiated) behavior correlates with higher SPIs (infrequent EODs) after 5000ms. **(d) Distribution of SPIs when the agent has high energy vs. low energy**, from 30 episodes of 1500 timesteps each. When the agent's energy is high, its discharge patterns show a wide range of SPIs, including high SPIs (infrequent EODs). When the agent's energy (satiety) is low, its SPIs are low because it is actively foraging with more frequent EOD discharges. Each SPI incurs a metabolic cost, so it is notable that the low-energy agent pulses frequently. This indicates that the low-energy agent prioritizes finding food to avoid starvation, rather than conserving energy. **(e) Distribution of agent energy levels and linear velocities** (in cm/timestep) across 30 episodes of 1500 timesteps each. The agent tends to maintain its energy at a "set point" close to full. Above this set point, the agent is penalized for overeating. The agent's linear velocity is bimodal (not swimming vs. swimming vigorously). High energy levels (high satiety) correlate with low velocity, and conversely, low energy levels correlate with high velocity. The agent's energy level appears to influence its locomotion strategy. When the agent's energy level (satiety) is high, it does not need to eat more food and swims slowly. High velocity often corresponds to an agent motivated to eat more food and gain energy. **(f) Feature importance from a 100-tree random forest predicting EOD rate**. Agent energy, followed by proximity to food, is the most important predictor of EOD rate. **(g) Principal component analysis (PCA) of the RNN's hidden states**, from 30 episodes of 3000 timesteps each. The hidden state output by the RNN at each timestep can be interpreted as a low-dimensional "summary" of the agent's belief about the state of the environment. We observe a transition along the 2^{nd} principal component between the hidden states corresponding to a low-energy agent vs. those of a high-energy agent. This indicates that energy may be a latent variable that plays an important role in determining the agent's actions.

References

billycorgan84. (2009). *Gnathonemus_petersii.* Wikimedia Commons. Retrieved from https://commons.wikimedia.org/wiki/File:Gnathonemus_petersii.jpg

Haesemeyer, M., Schier, A. F., & Engert, F. (2019). Convergent temperature representations in artificial and biological neural networks. *Neuron*, *103*(6), 1123–1134.

Ni, T., Eysenbach, B., & Salakhutdinov, R. (2021). Recurrent model-free rl is a strong baseline for many POMDPs. *arXiv preprint arXiv:2110.05038*.

Pedraja, F., & Sawtell, N. B. (2024). Collective sensing in electric fish. *Nature*, 1–6.

Rajan, K., Harvey, C. D., & Tank, D. W. (2016). Recurrent network models of sequence generation and memory. *Neuron*, *90*(1), 128–142.

Sawtell, N. B., Williams, A., & Bell, C. C. (2005). From sparks to spikes: information processing in the electrosensory systems of fish. *Current opinion in neurobiology*, *15*(4), 437–443.

Singh, S. H., van Breugel, F., Rao, R. P., & Brunton, B. W. (2023). Emergent behaviour and neural dynamics in artificial agents tracking odour plumes. *Nature Machine Intelligence*, *5*(1), 58–70.

von der Emde, G. (1992). Electrolocation of capacitive objects in four species of pulse-type weakly electric fish: Ii. electric signalling behaviour. *Ethology*, *92*(3), 177–192.

Von der Emde, G. (1999). Active electrolocation of objects in weakly electric fish. *Journal of experimental biology*, *202*(10), 1205–1215.