Neuronal shaping in a Co-Adaptive Brain-Machine Interface

Babak Mahmoudi¹, Jack DiGiovanna¹, Jose C. Principe², Justin C. Sanchez³

¹Department of Biomedical Engineering, ²Department of Electrical and Computer Engineering, ³Department of Pediatrics, Neurology University of Florida

A new framework for studying causation between biological networks and computational models has been developed using a motor Brain-Machine Interface (BMI) system based on Reinforcement Learning (RL). The closed-loop RLBMI shown in Fig. 1 is a framework to test theories of goal-based learning and decision making through experience using a robotic arm in 3-D space during a reaching task. Here, the interaction between an agent and user’s brain occurs through the generation of a sequence brain states that are mapped by the agent to a series of actions of a robotic arm. Using states, actions, and rewards, the agent and user must learn to co-adapt with each other to maximize the earned reward.

In this paradigm, we have trained 3 rats in a two-target choice task. To obtain a water reward, the rats were required to use a robotic arm to press one of two levers (left and right positions), cued by an LED light. The rats were bilaterally implanted with two electrode arrays in the forelimb region of primary motor cortex (16 electrodes in each M1). During brain-control of the robotic arm, single unit activity of cortical cells was chronically recorded and their firing rates (100ms bins) were used as environmental state in RLBMI. In this architecture, the agent is the robot controller and at each time step, the agent maneuvers by evaluating its environmental state and selecting one of the 26 actions in the 3-D space. Initially, both the rat and agent are naïve to the solution of the task. We have used a two layer neural network (MLP) for Value Function Estimation (VFE). In the VFE structure, trained with temporal-difference learning, the action corresponding to the output node with maximum value was selected for robot control. To test the neural response to the co-adaptation, the shaping of complex behaviors was enabled with an adjustable threshold which sets the necessary target proximity to earn reward. This threshold was iteratively adjusted from close to the robot starting position to far away where the probability of randomly intersecting the target was 25% - 9%.

Using this architecture, we demonstrate that the three animals were able to maintain a reward performance 460%, 517%, and 515% over chance despite the increasing task difficulty. Here we study two adaptive elements of the architecture the agent model and the user’s neuromodulation. First, in the agent’s VFE network we observed the weight tracks to be smoothly varying (no discontinuities) over multiple days indicating that past experience was being maintained by the agent for each difficulty level. Excited by the neuronal firings, the input layer of the MLP showed decreasing variability as task performance increased. Second, for the user’s neuromodulation, 60% of the neurons had a decrease, 30% had an increase and 10% had no significant change in the mean firing rate as a function of the difficulty level when compared to the first session. Interestingly, of the neurons with a decrease in firing rate, 73% had an increase in their coefficient of variation (CV) of firing and the rest had no significant change in their CV. 86% of the neurons which had increase in their mean firing had no significant change in their CV. The remaining 10% of the neurons that had no significant change in their mean firing rate also did not show a significant change in their CV. All metrics were tested for significance using ANOVA at 95%. Based on the results, co-adaptation does not primarily occur as a general up-regulation in neuronal firing but as an increase in temporally specific neuromodulation of the ensemble related to subgoals of the complete reaching task.

Acknowledgements: This work was supported by NSF Grant #CNS-0540304.