BMC Neuroscience

Poster presentation

Is self-control a learned strategy employed by a reward maximizing brain?

Aristodemos Cleanthous* and Chris Christodoulou

Address: Department of Computer Science, University of Cyprus, Nicosia, 1678, Cyprus

Email: Aristodemos Cleanthous* - aris@cs.ucy.ac.cy

* Corresponding author

This abstract is available from: http://www.biomedcentral.com/1471-2202/10/S1/P14

Self-control can be defined as choosing a large delayed reward over a small immediate reward [1]. Brain-imaging studies [2] have shown that such behaviors result from competition between neural systems demonstrating that two separate systems are involved in such decisions. In particular, parts of the limbic system are preferentially activated by decisions involving instant rewards whereas regions of the prefrontal cortex are engaged uniformly by intertemporal choices irrespective of delay [2]. Moreover, the subjects' choice was directly linked to the relative activation of the two systems [2]. As Kavka [3] suggests, it is possible that such inner conflicts are resolved as if they were a result of strategic interaction among rational subagents.

A computational model of interpersonal conflict is proposed where we implement two spiking neural networks as two players, learning simultaneously but independently, competing in the Iterated Prisoner's Dilemma (IPD) game. An interpretation of the IPD is that it demonstrates interpersonal conflict [3] where the Cooperate-Cooperate (*CC*) outcome corresponds to the behavior of self-control. The outcome of each round of the game is taken according to the relative output activation. The purpose of the system is to learn how to exhibit self-control through biologically plausible reinforcement learning. To the best of our knowledge, our work implements, for the first time, a game theoretical view of self-control with a computational system that learns through biologically plausible algorithms.

Learning in our system links behavior to the synaptic level by reinforcing stochastic synaptic transmission [4]. Results show that the system managed to maximize reward by establishing a strong self-controlled behavior, reflected by a strong *CC* outcome [5]. It is noted that the self-control outcome not only persisted during the final rounds of the games, but it also did not change after the 100th round due to the system's dynamics that were evolved by that point in time in such a way to consistently produce the self-control outcome. This reveals that after a certain point the networks learned that is for their own benefit to compromise in order to maximize their long-term reward. Preliminary results suggest that the system's performance, especially its adaptability, is further enhanced when reinforcement learning through modulated Spike-Timing-Depended Plasticity [6,7] is integrated into the system. Overall, our results indicate that self-control is a learned strategy employed by a reward maximizing brain in the presence of competing neural systems that results to the regulated activation of the respective systems.

Acknowledgements

References

1. Rachlin H: *The Science of Self-Control* Cambridge, MA: Harvard University Press; 2000.

2. McClure SM, Laibson DI, Loewenstein G, Cohen JD: **Separate neural systems value immediate and delayed monetary rewards.** *Science* 2004, **306:**503-507.
3. Kavka G: **Is individual choice less problematic than collective choice?** *Economics and Philosophy* 1991, **7:**143-165.
4. Seung HS: **Learning in spiking neural networks by reinforcement of synaptic transmission.** *Neuron* 2003, **40:**1063-1073.
5. Christodoulou C, Banfield G, Cleanthous A: **Self-control with spiking and non-spiking neural networks playing games.** *Journal of Physiology (Paris)* in press.
6. Florian RV: **Reinforcement learning through modulation of spike-timing dependent synaptic plasticity.** *Neural Computation* 2007, **19:**1468-1502.
7. Legenstein R, Pecevski D, Maass W: **A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback.** *PLoS Computational Biology* 2008, **4:**e1000180.