Neuromorphic reinforcement learning

Julie Dethier, Damien Ernst, and Rodolphe Sepulchre
Systems and Modeling Research Unit, ULg
10 Grande Traverse, B-4000 Liege, Belgium
Emails: [jdethier, dernst, R.Sepulchre]@ulg.ac.be

1 Neuromorphic engineering

Living organisms are able to successfully perform challenging tasks such as perception, classification, association, and control. In hope for similar successes in artificial systems, *neuromorphic engineering* uses neurophysiological models of perception and information processing in biological systems to emulate their functions but also resemble their structure [1]. In this abstract, we focus on the basal ganglia (BG), brain region in control of primitive functions of the nervous system, and specifically on their involvement in action selection and reinforcement learning (RL). We hypothesize that neuromorphic-inspired systems will greatly benefit the RL community.

2 Computational architecture of the basal ganglia

The BG are a group of interconnected subcortical nuclei that participate in cortical- and sub-cortical loops. These loops are topographically organized in relatively discrete channels that loop back, via appropriate thalamic relays, to the same area of cortex (e.g. limbic, associative, sensorimotor) from which they originated [2]. Two essential functions of the BG are action selection and RL; we investigate how theses functions can be morphed and engineeringly exploited.

2.1 Action selection

Parallel processing functional systems that compete for behavioral expression loop through the BG, conveying phasic excitatory signals—"bids" for selection—to the input nuclei [2]. Through comparison of input magnitudes (competing bids), the tonic inhibitory output is withdrawn from "selected" channels—disinhibition via the direct pathway of thalamocortical targets—and maintained or increased on "non-selected" channels—inhibition via the indirect pathways to suppress unwanted actions [3]. This action selection model can be exploited in Cognitive Pattern Generators, by analogy to the motor system's Central Pattern Generators, rhythm generators that operate to organize cognition [4].

Integration of these rhythm generators in Reservoir Computing (RC) models could generate powerful neuromorphic processing systems. RC, emulating information processing in the cortex, relies on a fully connected one-hidden-layer recurrent neural network, the reservoir, with dynamics at the "edge of chaos" and with only trainable weights in the connections from hidden nodes to the multiple outputs [5]. This simplicity of training comes with challenges: creating a rich enough reservoir, particularly if many dynamics systems employ its different outputs with different sets of weights [5]. Cognitive Pattern Generators could select specific system dynamics for a set of desired outputs.

2.2 Reinforcement learning

The BG play also a critical role in reward and RL circuits. Phasic firing in dopamine (DA) neurons in the ventral tegmental area (VTA)—BG region providing important modulatory signals to other BG nuclei and external structures—complies with a *reward prediction error* signal of contemporary learning theories, e.g. in temporal difference (TD) learning [6]. One suggestion for biological RL is DA modulation of cortico-striatal synaptic plasticity [8]. Exploiting this reward-modulated plasticity could improve RC effectivity: the internal dynamics can autonomously tune themselves to the dynamic regime which is optimal for a given task [7]. This mechanism could also explain cognitive functions, e.g. conditioning and working memory, and dysfunctions, e.g. Parkinson's and schizophrenia [8].

3 Basal ganglia model

The first step in this learning-oriented neuromorphic engineering is the modeling of the BG and their parallel processing loops, a subject of ongoing research. Particular interest lies in phasic firing in DA neurons and its role in plasticity.

References

[1] C. A. Mead, *P IEEE*, **78(10)**, 1990: 1629–39.

[2] F. A. Middleton and P. L. Strick, *Brain Res Rev*, **31**, 2000: 236–50.

[3] T. J. Prescott, F. M. M. Gonzalez, K. Gurney, M. D. Humphries, and P. A. Redgrave, *Neural Networks*, **19**, 2006: 31–61.

[4] A. M. Graybiel, *Schizophr bull*, **23(3)**, 1997: 459–69.

[5] D. Prokhorov, *IEEE IJCNN*, Montreal, Canada, 2005.

[6] W. Schultz, *Annu Rev Psychol*, **57**, 2006: 87–115.

[7] B. Schrauwen, M. Wardermann, D. Verstraeten, J. J. Steil, and D. Stroobandt, *Neurocomputing*, **71(7-9)**, 2008: 1159–71.

[8] R. D. Samson, M. J. Frank, and J. M. Fellous, *Cogn Neurodyn*, **4(2)**, 2010: 91–105.