Real Time Learning of Rearranging Associations on a Humanoid Robot

Emeç Erçelik, Neslihan Serap Şengör
Electronics and Communication Engineering Department, Istanbul Technical University, {ercelike,sengorn}@itu.edu.tr
www.simmag.itu.edu.tr

 Bernstein Conference 2014, Göttingen

Abstract
In this study, we focused on modeling a temporal sequence task performed by macaque monkeys on a humanoid robot. In the task, it expected to match a stimulus with an appropriate action [1] and to rearrange the association built up. We utilized a computational model of Cortex-Basal ganglia-Thalamus loop to select appropriate action and reinforcement learning approach to realize updating model parameters (Wc & Wr) with temporal difference method[2,3].

Humanoid robot, Darwin-Op is used for implementation. The computational model is constructed using Spiking Neural Network for the Cortex part and using dynamical system mass model for Basal Ganglia-Thalamus part. This study is aimed to be a step to close the gap between the neuron based approaches and behavioral approaches used to realize learning in robotic applications. The task is structured upon matching a color with a desired action and rearranging the built up association to select a new action using reward.

The Task
As shown in the left diagram, when a color is presented to robot, it tries to decide on an action, at first randomly, and gets reward or not according to its decision. The reward gained, while the robot was expecting nothing as a result of its choice, causes an expectation error. This urges robot to reinforce the desired stimulus-action pair. Once robot is rewarded due to a right choice, expectation error updates the parameters of BTC model in charge of action selection. When update is completed, robot manages to select the desired action to a given stimulus. In case, when an expected reward is not given to the selected action associated with a stimulus, this also causes an expectation error and urges robot to rearrange the previously built association between a stimulus and action.

The Model
The model consists of two parts, one is Cortex that is modeled by using SNN[4] and the other is Basal ganglia-Thalamus part that is modeled by using dynamical system approach [5]. The BTC model is coded in Python and the Cortex is simulated in NEST[6]. There are three channels for a different action. And the winning channel in Cortex is considered as the selected action.

The camera information is get from the C++ coded part and sent to the model to get action decision. The output of Cortex is modeled by using firing rates and this information sent to the Basal ganglia-Thalamus loop in a normalized form to manage the calculations. At the end of this calculation the output of Ctx is sent with Striatum information of model [to be use in RL] to update the model parameters with using TDL in C++ coded part.

Acknowledgements
* Thanks to Berat DeveciDurduran for his valuable discussions.
* This work has been supported by Tubitak Project 1112E264.

References