Real Time Learning of Rearranging Associations on a Humanoid Robot

We considered a temporal sequence task performed by macaque monkeys where to match a stimulus with an appropriate movement[1] and rearranging the association is expected. We utilized Basal ganglia-Thalamus-Cortex(BTC) model to realize action selection and to select appropriate action, reinforcement learning rules are used to update BTC model parameters using temporal difference method. In [2,3], it is claimed that temporal difference method is biologically meaningful. This working memory task is realized on humanoid robot, Darwin-Op on which online training is carried out.

During task, robot learns to associate different stimulus with desired actions as summarized in figure. There are three different stimuli which are red, blue and yellow colours and three different actions. Robot recognizes the colours with the embedded camera and detected colour stimulus is the input of BTC model, where action selection is done. Robot is expected to decide on an action at first. Green colour is reward, and it is shown to robot if the selected action is the correct to be associated with the stimulus. At first, robot was expecting nothing and a given reward raises its expectation error. This urges robot to reinforce the desired stimulus-action pair. Once robot is rewarded due to a right choice, expectation error(ee) 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 ee and urges robot to rearrange the previously built association between a stimulus and action.

So, robot manages to rearrange its associations by changing parameters of BTC model, which directly effects action selection and is biologically meaningful. As a future work, it is aimed to investigate the biological correlation between the parameters of BTC model and role of dopamine on action selection.

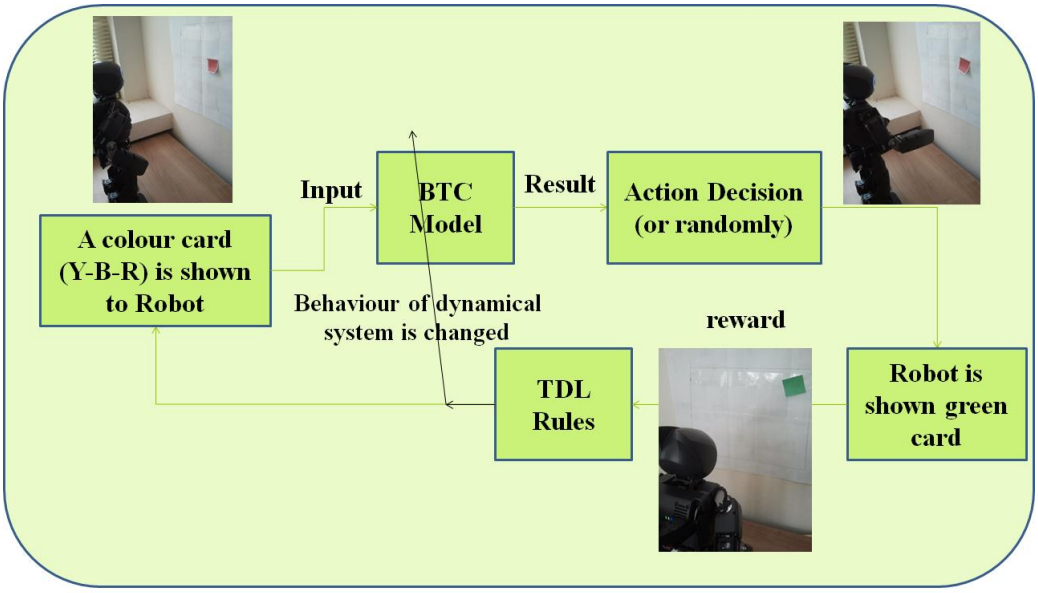

Figure 1: The scheme of implementing real time learning on humanoid robot

Acknowledgement

This work is supported by project TUBITAK 1001 no: 111E264.

References:

[1] Tanji J. & Shima, K., 1994. Role For Supplementary Motor Area Cells in Planning Several Movements Ahead. Nature **371**, 413-416.

[2] Schultz, W., Dayan, P. and Montague, P. (1997). A Neural Substrate of Prediction and Reward, Science,275, pp.1593-1599.

[3] Denizdurduran B., Sengor N.S., 2012. A Realization of Goal-Directed Behavior-Implementing a Robot Model Based on Cortico-Striato-Thalamic Circuits. ICAART 2012, Algarve, Portugal.

[4] Prescott, T. J., Montes-Gonzalez, F. M., Gurney, K., Humpries, M. D., Redgrave, P., 2006. A Robot Model of the Basal Ganglia: Behaviour and Intrinsic Processing. Neural Networks, 1-31.