A model of cerebellar automation of voluntary basal-ganglia control

Miika Pihlaja and Harri Valpola

HELSINKI UNIVERSITY OF TECHNOLOGY
Laboratory of Computational Engineering

Computational Neuroscience Group
www.lce.hut.fi/research/eas/compneuro/

Introduction

Cerebellar learning can be roughly categorized as supervised learning, while reinforcement learning is used as a model of basal-ganglia function [1]. Basal ganglia learn voluntary actions by trial and error [2], whereas cerebellum is among other things specialized in automation and fine tuning of motor control [3].

Cerebellum is often modelled as a feedback error learner, where the output signal of non-adaptive feedback controller is used as a teaching signal for the adaptive feed-forward controller [4]. Biologically this teaching signal can be thought of as a *reflex*.

- Learning new task requires a new handcrafted reflex signal every time
- Designing workable reflex signal becomes tedious with complex tasks

This can be circumvented by using a reinforcement learner to acquire a coarse version of the required feedback controller from one-dimensional reward signal [5].

Model architecture

The model is built on top of a classical actor-critic [6] architechture. However, our actor consists of two components:

- Policy in basal ganglia, which is teached according to the TD-error signal from the critic
- The cerebellar predictor, which gets the output of the basal ganglia as a teaching signal

The basal ganglia component of the actor learns a coarse policy according to the reward. The cerebellum component learns to predict these actions, and should eventually take over the control from basal ganglia.

- Value function is represented by a function approximator, which can consist of e.g. basis-function expansion followed by linear mapping, or MLP-net
- Basal ganglia component of the actor is a nonlinear mapping from state to a set of actions
- Cerebellar component is a supervised learner, which learns to predict the output of basal ganglia under uncertainty of the temporal delay in the system.

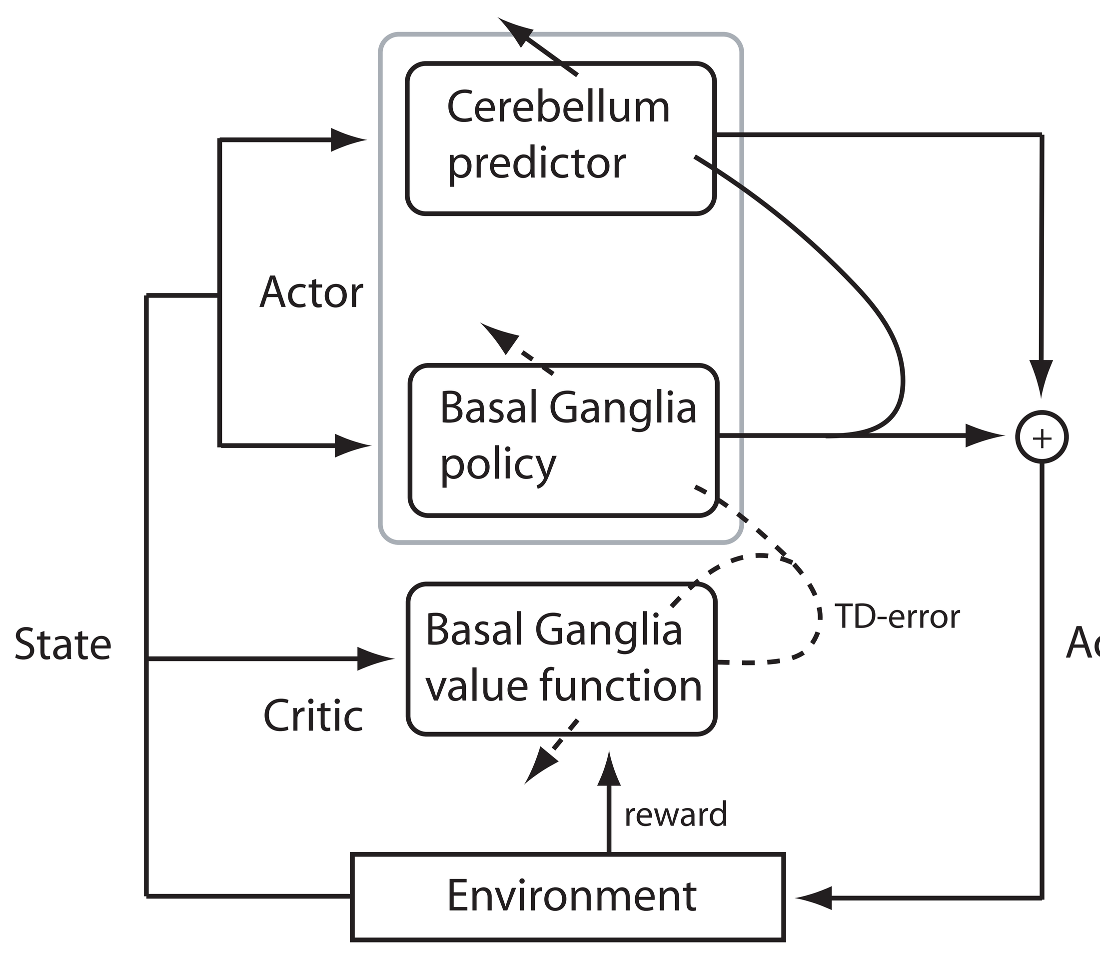


Figure1: The architecture of extended actor-critic algorithm

Neuroscience

Figure 2 illustrates a proposed mapping between our algorithm and a biological brain. Both basal ganglia and cerebellum receive information about the state of the world from cortical and sub-cortical structures. Value function is represented in the loops between striatum and VTA, while striatum together with the output structures of basal ganglia implement one part of the actor.

Cerebellar cortex together with deep nucleus (output structure of cerebellum) implement accurately timed mapping between inputs and outputs, whereas the inferior olive is responsible for giving the teaching signal. This forms the other part of the actor. Our model assumes that the output of basal ganglia is conveyed as an input to inferior olive e.g. through cortical pathways (dashed line in Figure 2).

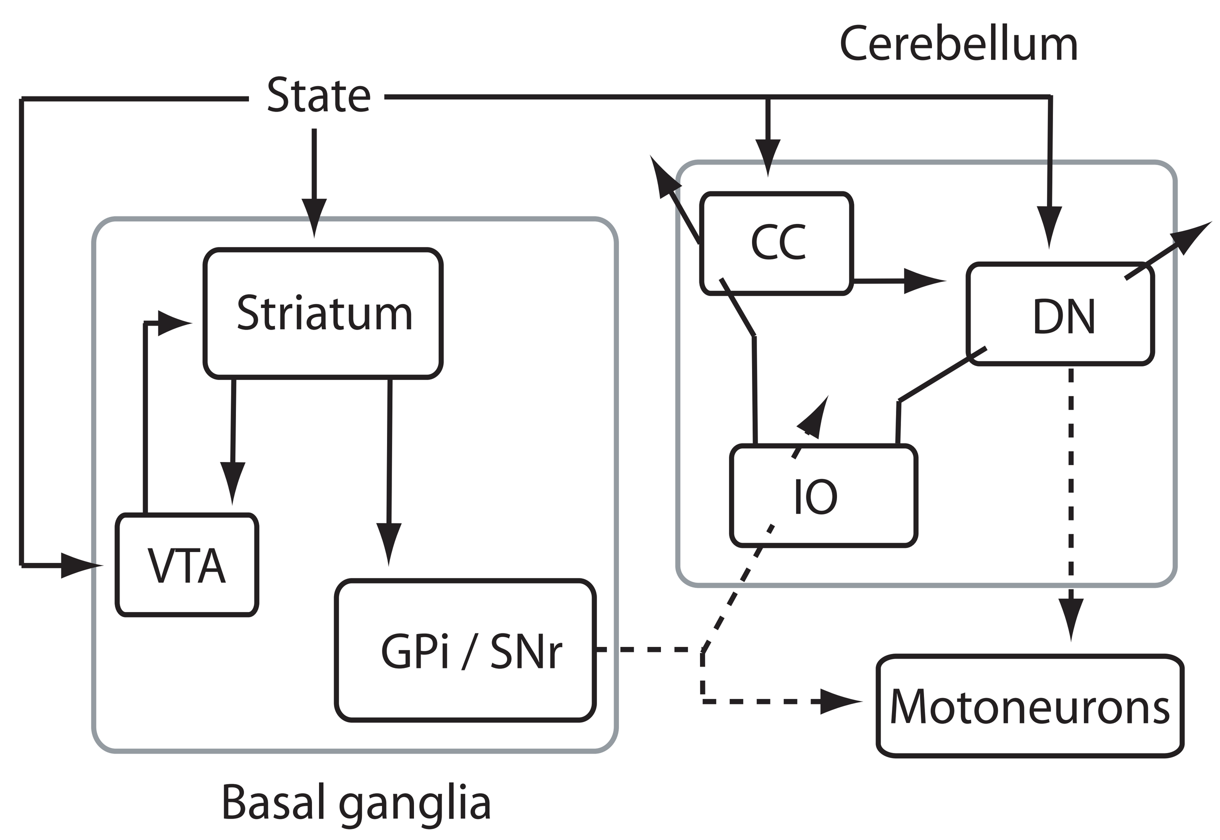
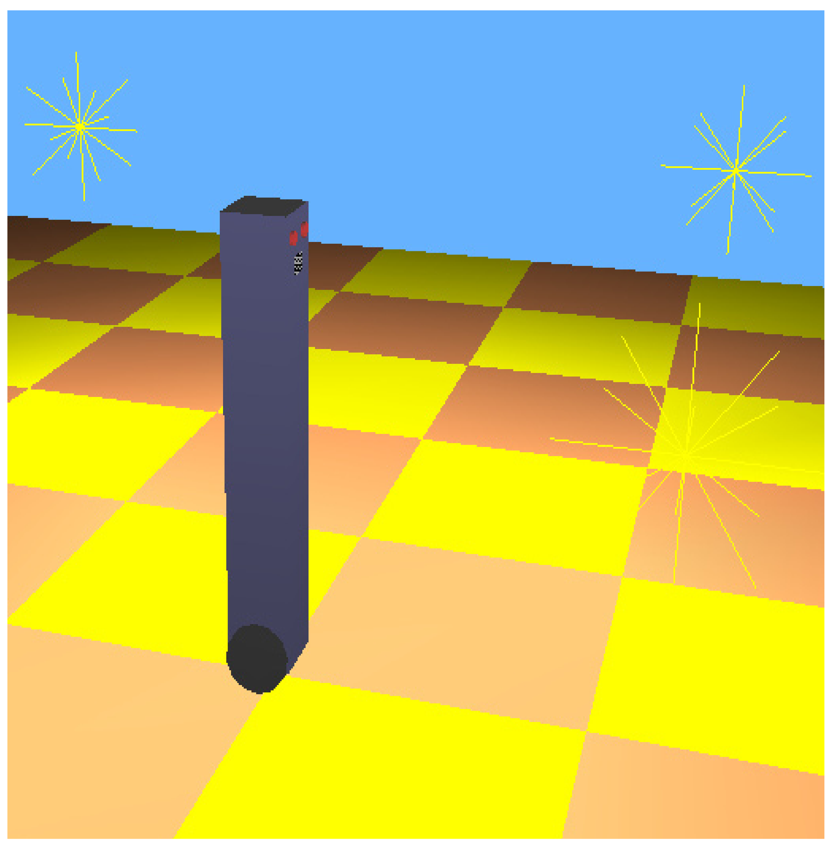


Figure 2: Proposed mapping of the algortihm to the biological brain. Ventral Tegmental Area (VTA), internal segment of globus pallidus (GPi), Substantia nigra pars reticulata (SNr), cerebellar cortex (CC), inferior olive (IO) and deep nucleus (DN).

Simulations

- Inverted pendulum with a cart, negative reward is proportional to the angle of the pole and it's displacement from the center of the arena.
- Simulation runs on Webots, which is linked to MATLAB for the implementation of the algorithm
- Still work in progress, initial results are promising



Main challenges

- Finding suitable representation for value function
- Finding a way to explore more efficient policies without causing catastrophical events (falling of the pole)

References

[1] K. Doya (1999), Neural Networks. 12(7-8):961-974.
[2] A. Graybiel (1995), Curr. Opin. Neurobiol. 5(6):733-41.
[3] M. Ito. (2002), Annals of the New York Academy of Sciences 978:273-288.
[4] M. Kawato (1997), Biol. Cybern. 57(3):169-85.
[5] H. Kambara et al. (2004), Proc. of the 26th Conference of the IEEE EMBS: 486-489.
[6] R. Sutton & A. Barto (1998), Reinforcement Learning: An Introduction.