Toy Model Report

24 May 2019

1 Previous stuff

1.1 Objective

To show the utility of structural plasticity by using the features of the song bird vocal learning circuitry.

1.2 Plan

Song birds have a direct motor pathway (SMP) and a parallel pathway (AFP) that helps with learning. Based on evidences from literature, we model the SMP using Hebbian learning, and the AFP using reinforcement learning. The influence of AFP is eventually consolidated within the SMP, rendering it capable of producing song independently. Also, there is evidence that the AFP drives motor learning prior to the SMP (in a crude sense). We attempt to provide a proof of concept that 1. there is a use of having 2 parallel pathways in learning, and 2. there is an advantage in having a head start for the AFP, by using the paradigm of arm exploration.

1.3 Songbird model

Initially, we model the two parallel pathways as shown in Fig1.

With this model, 1. Mimicking AFP was causing the SMP to learn song, and become independent eventually. 2. There was an advantage in RL having a headstart. However, AFP was able to produce song independently. [Fig 2]

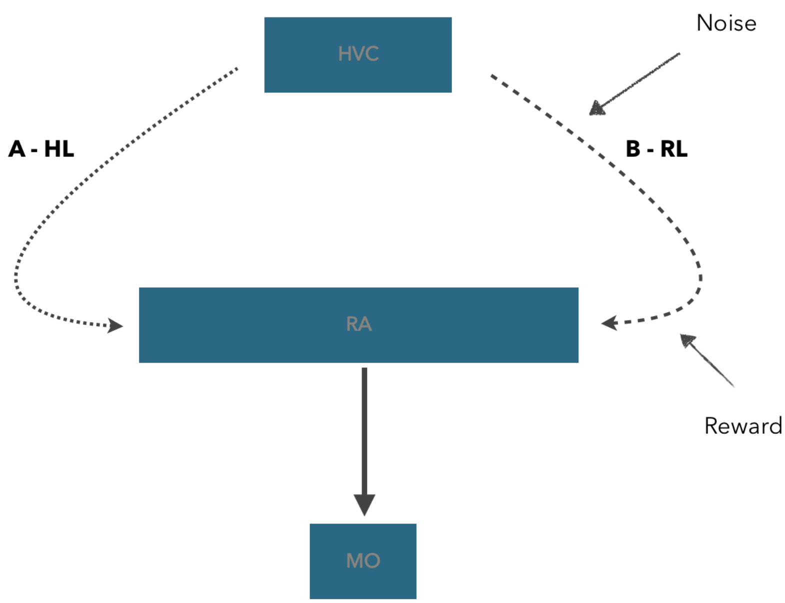


Figure 1: Schematic of model with SMP and AFP.

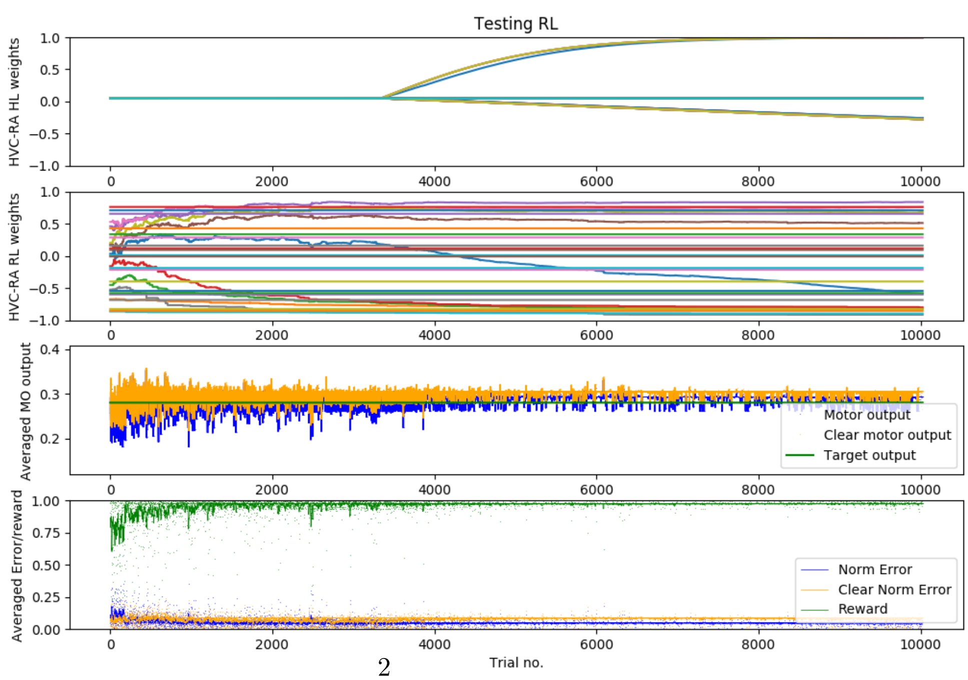


Figure 2: Model with SMP mimicking AFP. AFP is able to produce song independently, too.

To address this, we tried to limit the RL weights, to allow only local exploration. This led to 1. SMP's presence being necessary for song production. 2. AFP was not producing song independently. But, no advantage was observed when RL was given a headstart. [Fig 3]

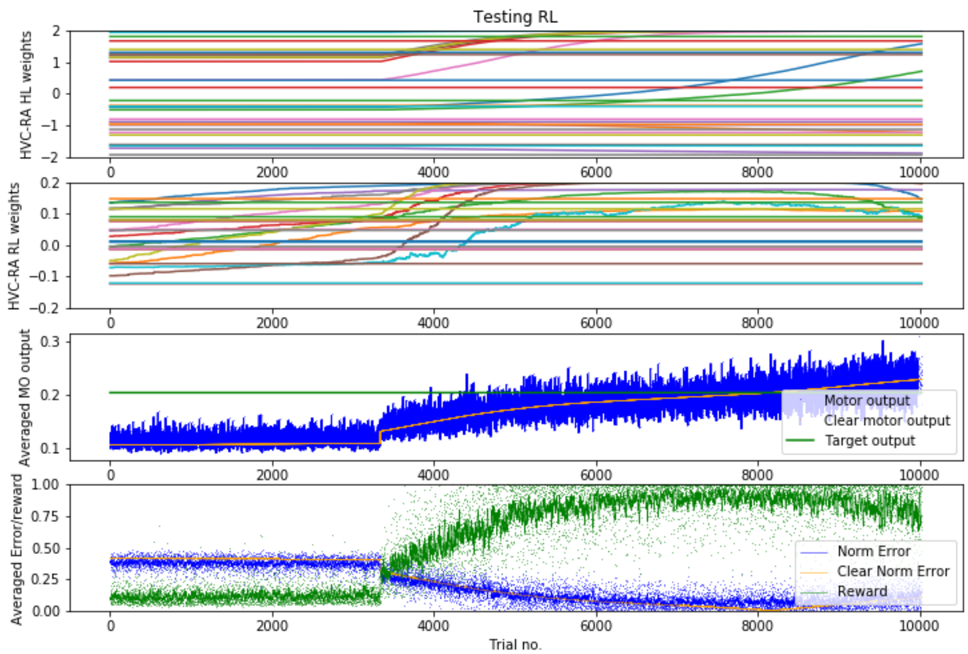


Figure 3: Model with limited AFP exploration.

To address this, we further look into this local motor exploration aiding SMP, we use the arm exploration paradigm. Task: A multi-segmented arm has the range to move in a circle (with its total arm length as radius), and is given a target to reach. We start with random exploration in the motor space. [Fig 4] The configuration of angles (between the arm segments) is perturbed is chosen. The reward is computed by comparing the end point of the arm with the target, and updating the angles accordingly.

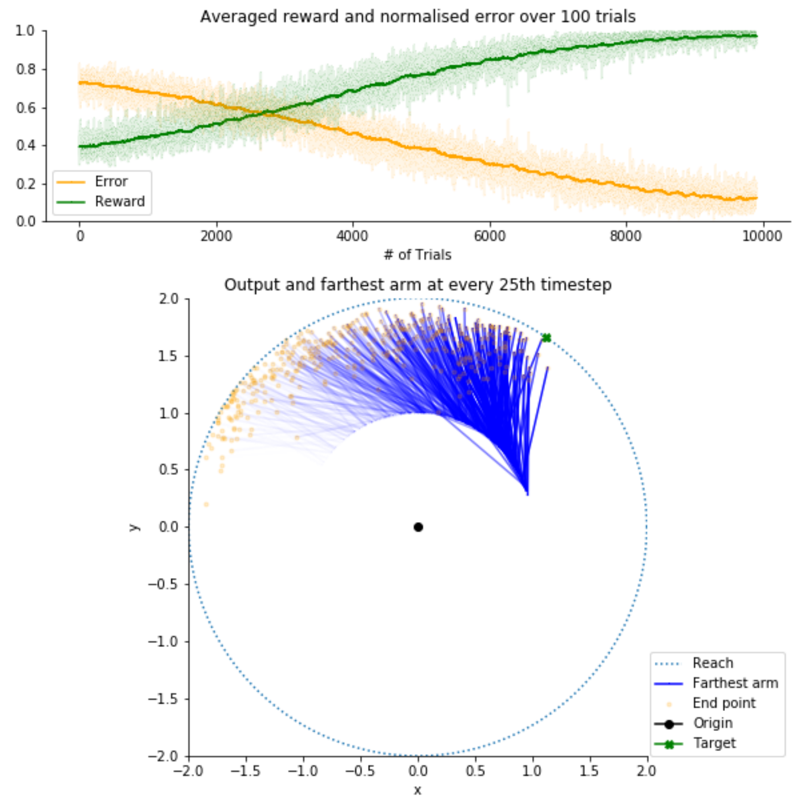
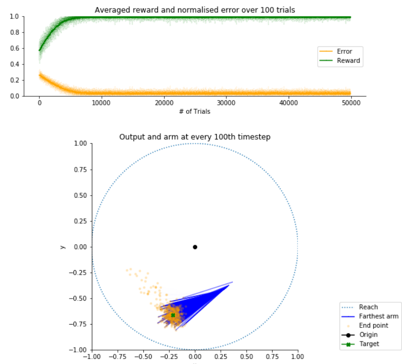
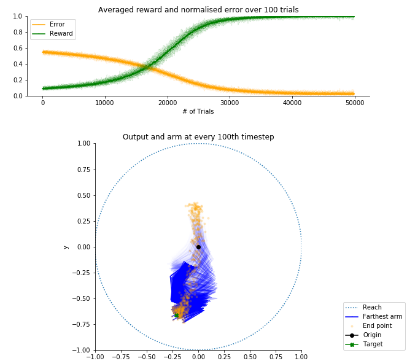


Figure 4: Arm model with random exploration.

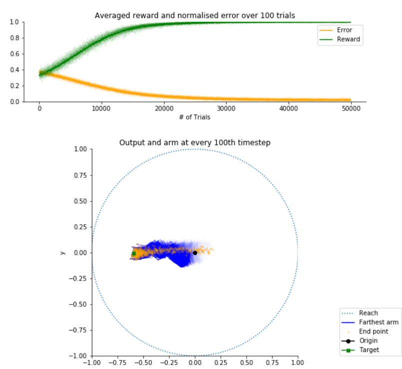
While, this leads to the target being reached eventually, in the case of a few no. of segments (n = 2), this is not the case as the no. of segments increases. This random exploration leads to higher exploration of certain regions of the available range (sensory space), more than others (typically centred around the origin). This is due to the fact that uniform exploration in the motor space, doesn't translate into uniform exploration in the sensory space, as the transformation between them is non-linear. Multiple arm configurations can result in the same endpoint. [Fig 5.]



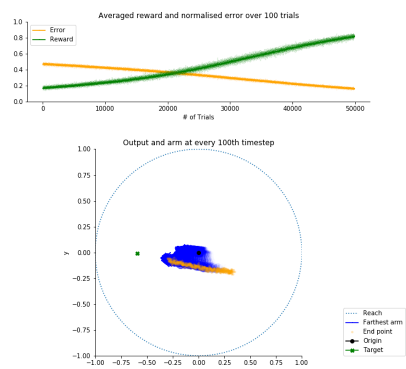
(a) 2 segments



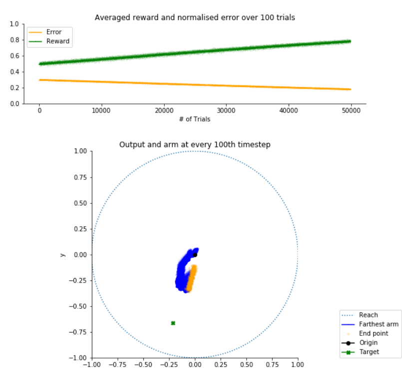
(b) 5 segments



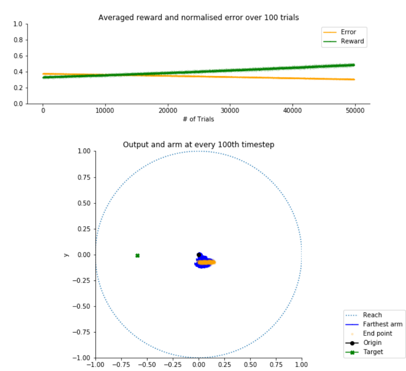
(c) 10 segments



(d) 25 segments



(e) 50 segments



(f) 100 segments

Figure 5: Exploration area reduces with an increase in no. of arm segments.

1.4 Arm exploration

We check if exploration in the sensory space resolves this. [**To be done.**]

2 Model to test now

Now, what I get from Arthur's suggestion of having motor exploration based on activity profiles.

2.1 Input

HVC - syllable encodings; governs which syllable will be produced.

2.2 Assumptions

For an abridged version for preliminary trial: 1 syllable (target), and 1 segment in arm. ($\implies$ we won't need an input for the 1 syllable case).

2.3 Workings

There is an activity function (Gaussian) for both the AFP and SMP, with an attenuating factor.

$$Act_{RL} = \frac{1}{\rho_{RL}} N(\mu_{RL}, \sigma_{RL}) \tag{1}$$

$$Act_{HL} = \frac{1}{\rho_{HL}} N(\mu_{HL}, \sigma_{HL}) \tag{2}$$

Initialise the standard deviation σ_{RL} high, indicating the higher variability in AFP.
Initialise σ_{HL} low, indicating the specificity of SMP driven motor activity.
The attenuating factor ρ_{HL} decreases over time, indicating the increasing strength of SMP connections.
ρ_{RL} increases over time, indicating the decreasing influence of LMAN.

$$Act_{total} = Act_{RL} + Act_{HL} \tag{3}$$

Take weighted average,

$$\theta_t = \frac{\Sigma_0^{2\pi} \theta.Act_{total}}{\Sigma_0^{2\pi} Act_{total}} \tag{4}$$

The angle (motor command) produced is a weighted average of the sum of the 2 activity functions. The effect is the polar coordinate reached on the plane of the arm.

$$Output = [r \cos \theta_t, r \sin \theta_t]$$

$$R = f(Output)$$

A reward is computed as an exponential function of the normalised error, where the error is the Euclidean distance between the target and the output coordinates. The mean of the activity functions are updated as follows.

$$\Delta\mu_{HL} = \eta_1(\theta_t - \mu_{HL}) \tag{5}$$

Hebbian learning tries to mimic the output of the whole system.

$$\Delta\mu_{RL} = \eta_2(\theta_t - \mu_{RL})(R - R_{prev})$$

(6)

Reinforcement tries to move towards higher reward.

$$\eta 1 << \eta 2$$

$$\mu_{HL} = (\mu_{HL} + \Delta\mu_{HL})\% angle_{max} \tag{7}$$

$$\mu_{RL} = (\mu_{RL} + \Delta\mu_{RL})\% angle_{max} \tag{8}$$

$$\rho_{RL} + +$$

$$\rho_{HL} - -$$

2.4 Choices

- For the activity function, we can either have a change in the attenuating factor, or a change in the standard deviation. I have chosen to change the attenuating factor, as (from what I understood in the discussion) want to depict a reduction in the influence of the LMAN.

2.5 Issues

- Initially, when there is no SMP influence, there is no difference in the weighted average from one trial to another, as the change in reward is zero, and the produced angle and the RL activity mean are the same. Solution: Can external noise be used?
- Learning rule for RL. — Tested, works with external noise.

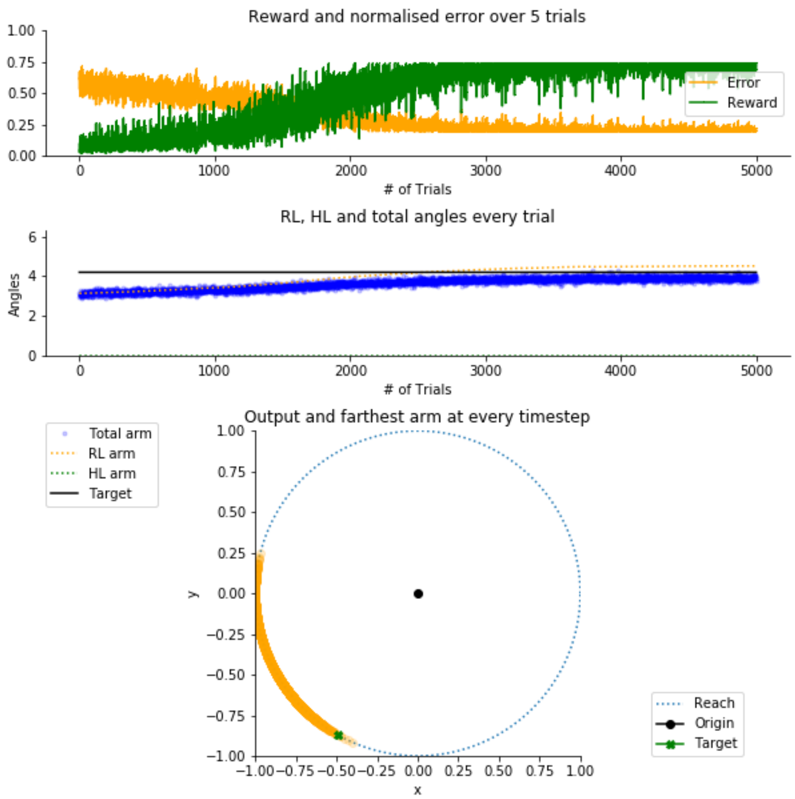


Figure 6: A starting attempt at the activity profile idea with external noise.