Action selection - Guthrie Model

Bhargav Teja Nallapu

September 7, 2015

List of Figures

1 Context

Basal Ganglia (BG) are known to be responsible for decision making and learning in a changing environment. *Gurthrie et al,* 2013 proposes a computational model for two-level action selection with two parallel functional loops each for cognitive and motor decision making. Each loop consists of a positive feedback direct pathway and a negative feedback hyper-direct pathway via the subthalmic nucleus to globus pallidus. In addition to the cognitive and motor ensembles, there are associative areas in cortex and striatum. The divergence and reconvergence of the direct pathway facilitates the information transfer between the cognitive and motor loops and hence providing a bias to the motor decision.

2 Materials and Methods

Task. The task used in the model to demonstrate action selection is a probabilistic learning task as described in Pasquereau et al. 2007. 4 target shapes are associated with different reward probabilities. A trial is a time period in which two of the four possible shapes are randomly chosen and presented (to the subject, *the model* in this case) at two random positions (of the possible positions - up, right, down and left). By the end of trial period, a choice is made by the *model* and the reward is presented (corresponding to the reward probability of the chosen shape). In a single independent trial, the cognitive decision (shape of the cue) and motor decision (direction of position) could be independent of each other.

Experiment setup. The first 500 ms of the trial is for settling of the model. Then the two random cues are presented. A trial is considered to be successful if a decision is made by the *model*, irrespective of the reward received. At any decision-making level of the *model*, each of the four cue shapes and each of the four motor movement directions is represented one unit (neuron) each. Thus in a given trial, when two cue shapes are presented at two different positions, two cognitive, two motor and two associative (wherever applicable) neurons are activated. The task is run for a session, a number of trials.

Learning. The *model* incorporates learning by taking the reward received in each trial into account forming a type of actor-critic network (Barto 1995). The learning changes the synaptic weights for the cortico-striatal gains over trials, particularily between the cognitive channels. The *model* design allows bidirectional information transfer between loops, a consequence of which was that, in early trials, a direction could be selected before a shape leading to a motor bias of cue selection(Fig. 1). However, after multiple trials , the *model* begins to consistently make cognitive decision before the motor decision in each trial and most frequently the motor decision towards the position of the more rewarding cue shape.

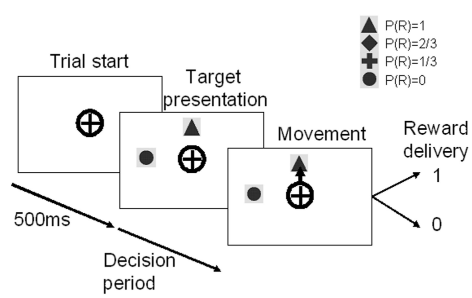


Figure 1: 2-armed bandit task in Guthrie model, as described in Pasquereau et al. 2007. The *model* has to make a cognitive decision about the shape and a motor direction decision, ideally towards the position of selected shape.

Accounting to the major characteristics of the Guthrie *model*, an elobarative and reproducible computational model[1] has been developed by *Meropi Topalidou*[2,3], *Arthur Leblois*[4], *Thomas boraud*[2], *Nicolas P. Rougier*[2,3]. This model reproduces the distinct activity of the chosen cue shape and the direction of chosen position (Fig. 2).

Neuronal dynamics. This model uses the same simplistic neuronal rate model as in Leblois et al. (2006) as used by the Guthrie *model* to focus on the network dynammics. Within each structure, each neuron (in each channel of any loop) is modeled as a single rate neuron with the equation:

$$\tau \frac{dm}{dt} = -m + I_s + I_{Ext} - T \tag{1}$$

decay time constant of the synaptic input τ, negative values of activation, m and threshold of the neuron T are set to respective constant values as per the Guthrie's *model*. I_{Ext} is an external input representing the sensory visual salience of the cue, which is unchanged thoughout the process of learning.

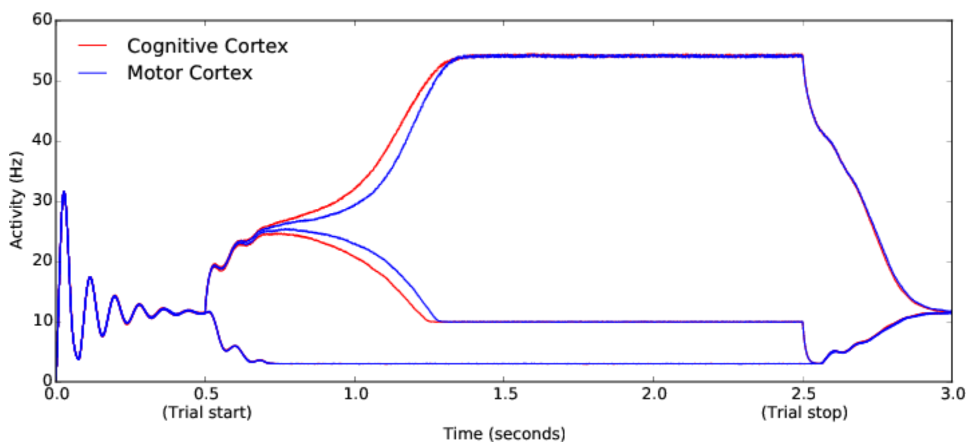


Figure 2: Activity (Hz) of all the channels in both cognitive and motor cortex varying with the trial time (seconds). The cognitive and motor activity corresponding to the cues which are not presented is constant(nearly zero). Of the 2 cues presented, the activity can be seen changing at a point (when decision is made). Also, it can be seen that for the cue whose activity increased (implying the cue has been chosen), cognitive activity (in red) rises before the motor activity (in blue) implying that the cognitive decision has been made before the motor decision.

When the *model* is presented over a session of many trials (120 in this case), the two cue shapes in each trial are randomly chosen from the four.

[1] http://www.labri.fr/perso/nrougier/downloads/ReproducibleScience.pdf
[2] Institute of Neurodegenerative Diseases - Bordeaux
[3] INRIA Bordeaux SUD-OUEST
[4] Laboratory of Neurophysics and Physiology Universite Paris Descartes

And it is made sure all the pairwise combinations of the shapes appear equal number of times throughout the session (6 combinations, 20 times each). Towards the last trials of the session, the model having learnt, the motor decision is biased towards the position of the cue shape which has been more rewarding from the beginning of the session. Many such independent sessions are run (in this case, 250). The performance of the model is measured by checking if the model made a motor position decision corresponding to the best rewarding cue shape of the two presented. This is measured as an average of each trial over different sessions. It is also reproduced that the mean performance of the *model* reaches close to 1 (Fig. 3), implying that after learning, the *model* most certainly selects the direction of the most rewarding cue in a given trial.

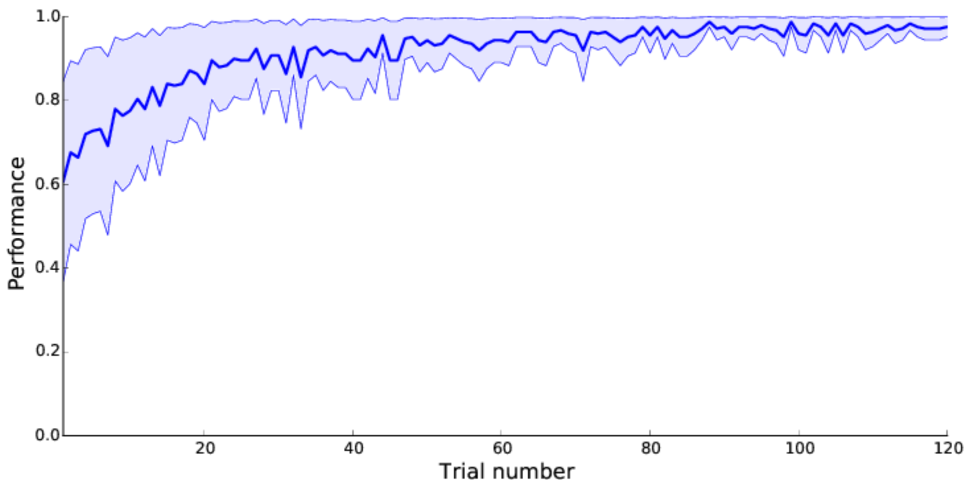


Figure 3: Performance of the model over various simulations.Each simulation consisting of 120 trials, is independent of one another. The learning that happens over 120 trials in a simulation is reset by resetting weights at the beginning of each simulation. A decision is considered correct if the motor decision made is in the direction in which the more rewarding cue is present. The performance is an average over all the simulations for each trial. It can seen that as the learning happens upto 120 trails, model tends to perform nearly accurate i.e, selecting the more rewarding cue almost every trial

3 Developments

3.1 Learning in cortico-striatal motor channels

As mentioned in *Learning, Materials and Methods*, the *model* learns only in the cognitive channels of the cortico-striatal areas, assuming the reward probability os associated to the cognitive aspect, i.e, the shape of the cue. The *model* is replicated, and learning is incorporated between the cortico-striatal motor channels. In every trial of the session, the weights of the cortico-striatal motor channels (position of the cue) are also adjusted besides that of cognitive channels upon the reward given. It can be observed that because of the learning, the synaptic weights are adjusted such that the weight of the channels associated to the more rewarding cue shapes is distintly more than the others (Fig.4).

Reduced performance. With the addition of motor channel learning, the *model* is run through 250 simulations, each with 120 trials. The learning rates , cortico-striatal gains and other parameters are set as per the original *model*. Besides the updated weights of motor channel, a reduced mean performance is observed when compared to the *model*'s mean performance in the absence of motor channel learning (Fig.5).

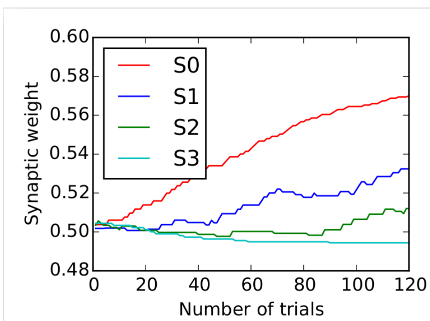
Faster motor decision. The original *model* with learning only in the cognitive channels, shows that motor decision time is consistently higher than the cognitive decision time in each trial, suggesting that the best shape is chosen and thus a motor decision is taken depending on the position of chosen shape. However, with the addition of learning in motor channel to the *model*, motor decision did not occur consistently after the cognitive decision implying that motor decision is hardly biased by the cognitive decision (chosen shape) (Fig. 6) and thereby affecting the overall performance.

3.2 Non-simultaneous cue presentation

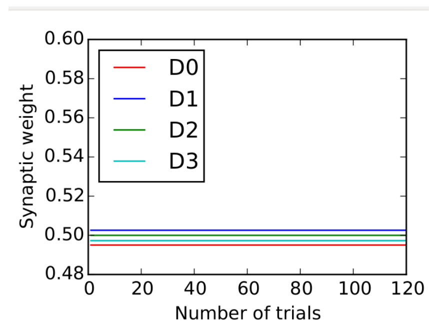
In the task presented in the original *model*, two randomly chosen cue shapes are presented simultaneously. And thus the *model* starts to make decision having both the cues already presented. However, if there is a certain delay between presenting both the cue shapes, it might affect the decision making. The *model* is allowed to learn for 120 trials with the usual task. Once the *model* has learnt the cortico-striatal cognitive channel weights, the *model* is tested with a trial in which only one cue shape is presented after the settling time. After a certain milliseconds of delay, the second shape is presented.

If the first presented cue shape is best rewarding compared to the upcoming one, it is expected that the *model* chooses the first cue and delay has no effect on the model. Hence the model is tested for the combination of cue presentations where first presented cue shape is less rewarding than the upcoming cue. For the each of the 6 possible combinations of cue presentations where the first presented cue shape could be less rewarding than the upcoming one, the *model* is tested for 120 trials. The performance is measured for each combination, the expectation being the *model* would select the best rewarding cue shape even if it is presented after a certain delay.

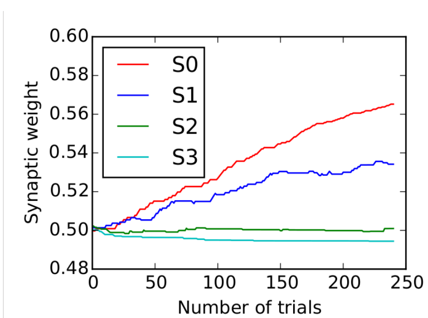
For smaller delays such as 10 ms or 20 ms, *model* is observed to select the best rewarding cue of the both, though it was presented after the delay. But as the delay increases, the performance of the *model* is observed to be reduced significantly. Interestingly, the delay after which the *model* couldn't



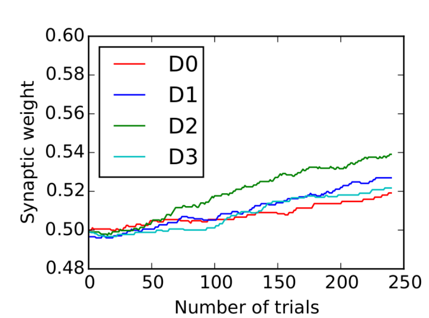
(a) Cognitive channel weights



(b) Motor channel weights



(c) Cognitive channel weights



(d) Motor channel weights

Figure 4: Synaptic cortico-striatal weights changing over trials. 4(a)&(b) as in the original *model* with learning only in cortico-striatal cognitive channels. Weights of channels are shown in different colors. S0,...S3 represent weights of channels associated with each of four cue shapes whereas D0,...D3 represent that of each position where cues could be shown. 4(a) shows that eventually after learning, the weight of one channel that of highly rewarding cue shape is distinctly higher than that of other cue shapes. As there is no learning in motor channels, the motor channel weights are unchanged as in 4(b). After modifying the *model* to add learning in cortico-striatal motor channels, 4(c) shows that the cognitive channels still learn as expected, increasing the weights as per the reward associated with the cue shape. At the same time, the motor channel weights are observed to be changed 4(d). Although reward given was not associated to the position of the cue presented, the *model* attributes certain positions to be rewarding by increasing motor channel weights.

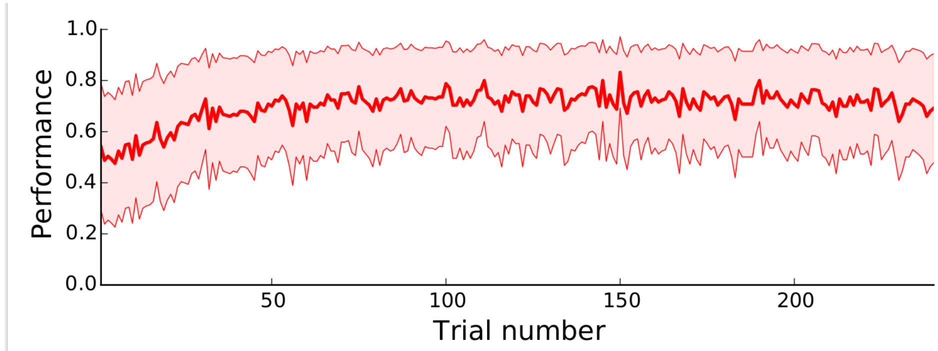


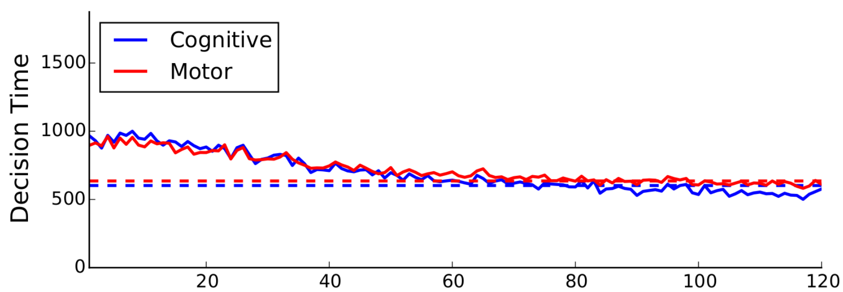
Figure 5: *Model* performance in the presence of motor channel learning. As it can be compared to Fig.3, the performance is found to be reduced when the cortico-striatal motor channels learn through the session.

select the best rewarding cue, is different for different combinations, depending on the best rewarding cue shape in the combination. The future work on the *model* includes making the *model* account for the delay between the cue presentations and still be able to choose the best rewarding cue.

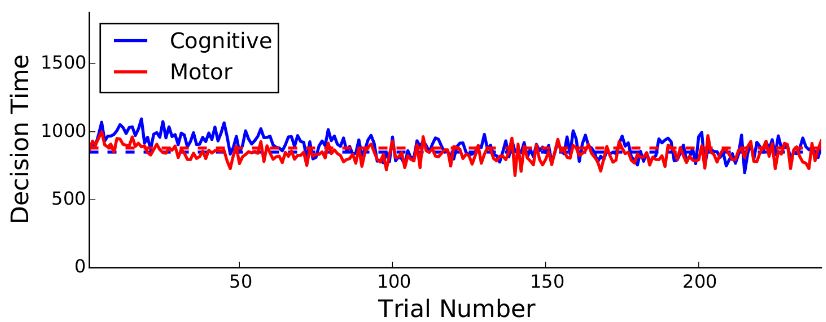
3.3 Visual salience of the cue

The *model* encodes the visual salience of the cue shape presented as I_{Ext} in the neuronal equation (Eq. 1). I_{Ext} is unchanged throughtout the learning process. After the *model* has learnt, change in such salience of the cue shape affected the performance of the *model*. In similar terms as *Section 3.2* , the *model* is allowed to learn for 120 trials. In a trial when two randomly chosen cue shapes are presented simultaneously, but the lesser rewarding cue has more salience than what the *model* has learnt, it is expected that the model still selects the best rewarding cue.

In each of the 6 pairwise cue combinations, the salience of the lesser rewarding cue shape is increased and each combination is presented for 120 trials. The performance of the *model* is measured, expecting the *model* still selects the best rewarding cue shape. However, the performance of the *model* is observed to be reduced when the lesser rewarding cue shape is presented with a certain increase in salience. Interestingly, the increase of salience after which the *model* performance decreases for each combination depends on the best rewarding cue shape in the combination.



(a) Original model



(b) Model with learning in motor channels

Figure 6: Decision times for each trial. The lines with dash represents the average decision time across all the trials. In the original model (a), when there is no learning in the motor channels, the motor decision time (in red) is more than the cognitive decision time (in blue), towards the ending trials. When the learning in motor channels is introduced (b), there is no such consistency of motor decision time observed, which implies motor decision (direction of position) was not necessarily biased by the cognitive decision (cue shape)

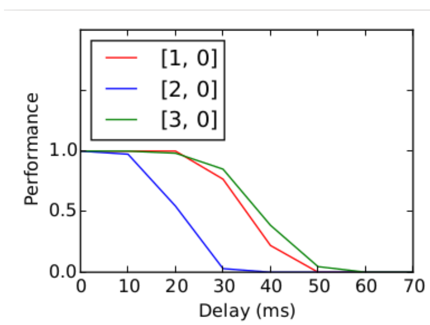
4 Progress

4.1 As on 21/08/2015

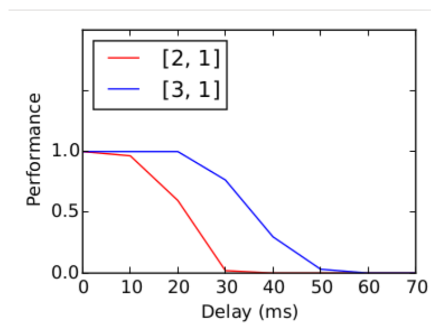
Scaling: The results of the developments as in the section 3 give some useful insights into various factors involving decision making. However, one of the major limitations of the *model* is the discrete representation of stimuli in the discussed areas. One neuron (channel) representing each cue shape stands too simplified a representation and limits us from analyzing cases like ambiguous cue shapes. With the idea of scaling up the model and have a number of neurons representing each stimulus, we decide to implement the exact model using *Distributed Asynchronous Numerical & Adaptive* *computing framework (DANA*[5]*)* and then scale up the number of neurons representing the cue shape inputs in important areas of cortex and striatum.

Having a population of neurons representing each cue shape input, the is analyzed to find an appropriate method. A guassian input at the center of each popualation for a particular cue shape is tried, to begin with. It'll be explored how the *model* translates this kind of stimuli representation to further levels in decision making.

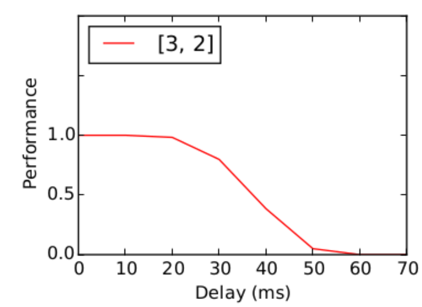
[5]http://dana.loria.fr/doc/index.html



(a) Cue shape appearing with delay - 0



(b) Cue shape appearing with delay - 1

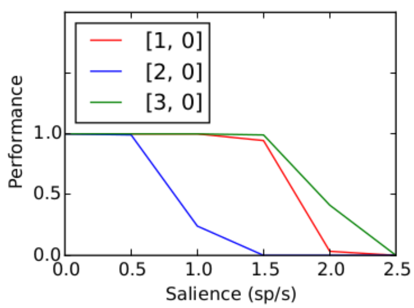


(c) Cue shape appearing with delay - 2

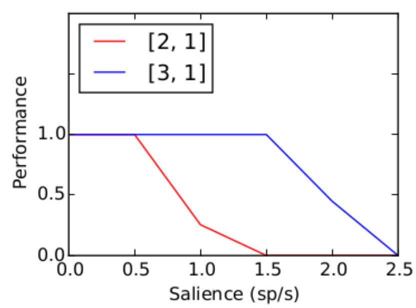
Figure 7: Delay in second cue presentation, performance of the *model*. [0,1,2,3] represent cue shapes in the decreasing order of reward probabilities [1.0,0.66,0.33,0] respectively. The *model* is allowed to learn in its original state. Then each combination of the following cue pairs are presented with the more rewarding cue appearing after a delay, for 120 trials and the performance is measured. (a) 3 combinations of cue shapes with cue shape - 0 appearing after a delay in all the cases. Performance starts to decrease as cue 0 is presented after a delay grerater than 10 ms. (b) 2 combinations with cue shape 1 appearing after a delay. Performance starts to decrease after the delay is greater than 20 ms. Similar observation with the cue shape 2 presented after a delay, (c).

DANA provides a sophisticated and simpler framework to project connections between structures of unequal populations. Also, being a python based framework, the internal representations used in *DANA* helps faster computation, particularily for dynamically changing neural populations, than numpy.

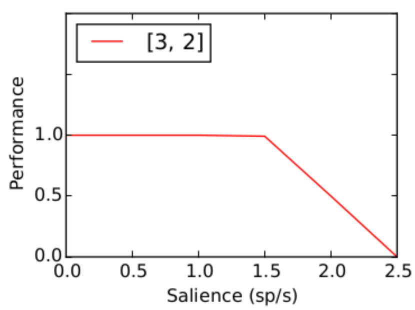
We have begun to convert the existing *model* to *DANA*. The idea is to retain the exact details of structures in the existing *model* and make sure



(a) Cue with lesser salience - 0



(b) Cue with lesser salience - 1



(c) Cue with lesser salience - 2

Figure 8: Cue shapes with different salience, performance of the *model*. [0,1,2,3] represent cue shapes in the decreasing order of reward probabilities [1.0,0.66,0.33,0] respectively. The *model* is allowed to learn in its original state. Then each combination of the following cue pairs are presented. In a trial, both the cues appear simultaneously, but the lesser rewarding cue is presented with increase in salience, for 120 trials and the performance is measured. (a) 3 combinations of cue shapes with cue shape - 0 with an increased salience to the lesser rewarding cue shape. Performance starts to decrease as the salience of the cue shape other than 0 is increased by more than 1 sp/s. (b) 2 combinations with cue shape 1 and the lesser rewarding cues with increased salience. Performance starts to decrease as the salience of the cue shape other than 1 is increased by more than 1.5 sp/s. Similar observation with the cue shape 3 presented with increased salience against cue shape 2 (c).

similar results are replicated; then expand the number of neurons in cortex and striatum. Owing to the knowledge that number of neurons in cortex is bigger than that in striatum, it is chosen to attempt the scaling up intially in cortex, beginning with 512 neurons in total while the projection remains to be onto 4 neurons in striatum. As the desired results are replicated, it would be fairly easy to scale up striatum.

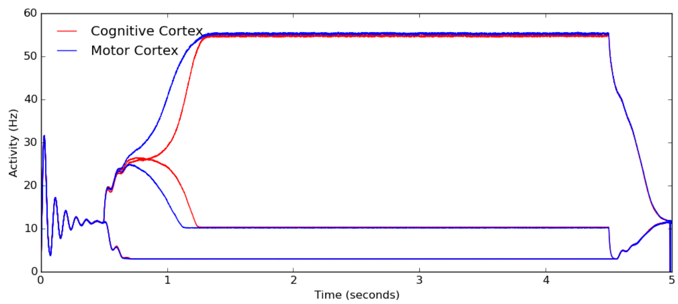
4.2 As on 04/09/2015

The model is converted to *DANA*, with structures connected through DANA connections. It is configurable to increase the population in cortex or striatum as needed. However, with increased population in cortex, the model needs to replicate the trial behaviour as exhibited in the Guthrie *model* (Fig. 2)

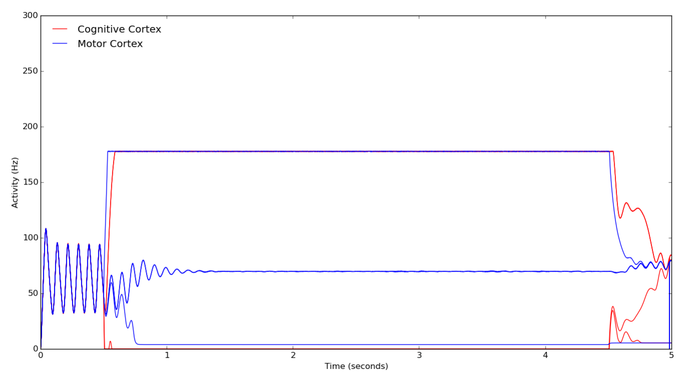
Input is represented by a gaussian, over the population representing each cue shape, with mean being at the center of the population. Unlike the simplistic case in the previous *model*, where a single neuron receives external current, for each cue shape, now the input is distributed across a population. Hence at the target structure, the gains need to be changed. The population in the cortex representing each cue shape (and each motor direction) is increased to 16 from 1 and that of striatum remaining 1.

The sigmoidal behaviour of the cortex(cognitive and motor) activity within the first 500 ms (settling time of the trial), which is a very significant feature of the Guthrie *model*, is reproduced. However the later stabilization of the activity representing both the cues, with higher activity implying selection of a choice, is not observed.

TBD - following week Identify suitable gains for the cortex to striatum projections, so that action selection can be observed, with a higher population in cortex to that in striatum.



(a) Cortex activity in a trial - 1 neuron per cue shape



(b) Cortex activity in a trial - 16 neurons per cue shape

Figure 9: Cortex activity in each trial. With the model implemented in *DANA*, and the population remaining same as the Guthrie *model*, the activity cab seen as identical to that of in the Guthrie *model* (a). When the the population is increased in cortex, the acitivity before the introduction of stimulus is identical. However, to obtain a selection of action, appropriate gains are to be identified.

References

[1] Martin Guthrie, Arthur Leblois, André Garenne, and Thomas Boraud. Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study. *Journal of neurophysiology*, 109(12):3025–3040, 2013.

[2] Arthur Leblois, Thomas Boraud, Wassilios Meissner, Hagai Bergman, and David Hansel. Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *The Journal of Neuroscience*, 26(13):3567–3583, 2006.

[3] Benjamin Pasquereau, Agnes Nadjar, David Arkadir, Erwan Bezard, Michel Goillandeau, Bernard Bioulac, Christian Eric Gross, and Thomas Boraud. Shaping of motor responses by incentive values through the basal ganglia. *The Journal of neuroscience*, 27(5):1176–1183, 2007.