

# Action selection - Guthrie Model

Bhargav Teja Nallapu

## 1 Context

This document refers to *model*<sup>1</sup> implemented by *Meropi Topalidou*<sup>2,3</sup>, *Arthur Leblois*<sup>4</sup>, *Thomas boraud*<sup>2</sup>, *Nicolas P. Rougier*<sup>2,3</sup>

The *model* in its original version models action selection using an N-armed bandit task. With reference to Guthrie et al. (2013), the *model* includes two distinct cognitive and motor loops to account for action selection at multiple levels. However the learning happens only in cognitive loop assuming that its the shape of the cue presented that matters to make a decision depending on the reward associated to it.

## 2 Scenarios

### 2.1 Motor learning

Here, we try to test, with motor learning in place, for different learning parameters, how the cognitive and motor weights (for directions) are learnt and how mean performance varies after various sessions.

---

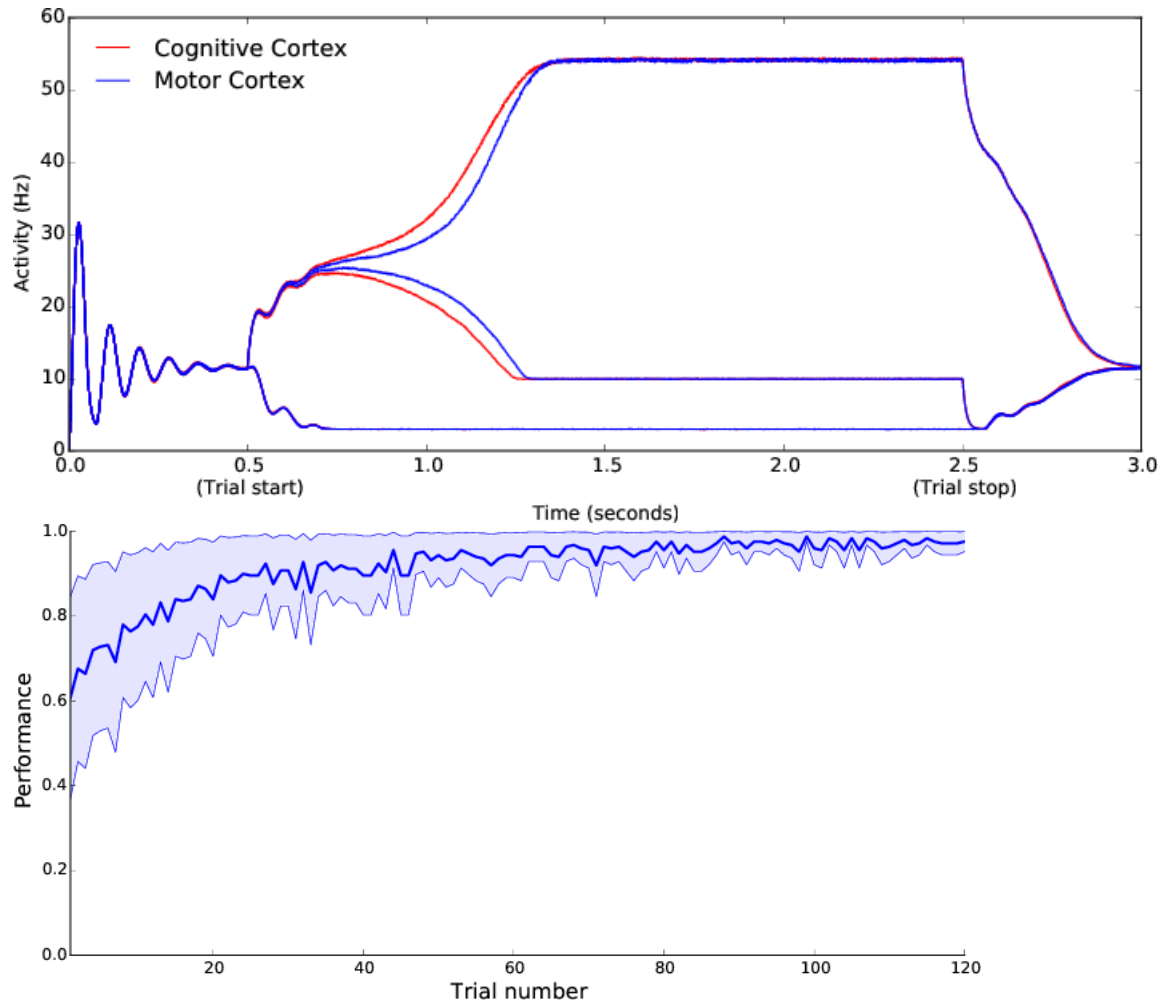
<sup>1</sup><http://www.labri.fr/perso/nrougier/downloads/ReproducibleScience.pdf>

<sup>2</sup>Institute of Neurodegenerative Diseases - Bordeaux

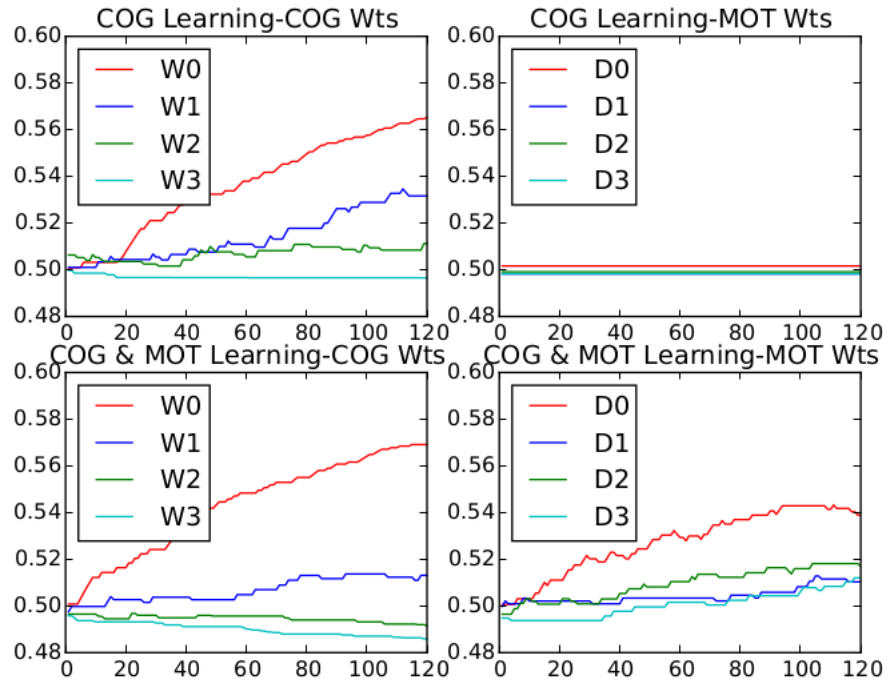
<sup>3</sup>INRIA Bordeaux SUD-OUEST

<sup>4</sup>Laboratory of Neurophysics and Physiology Universite Paris Descartes

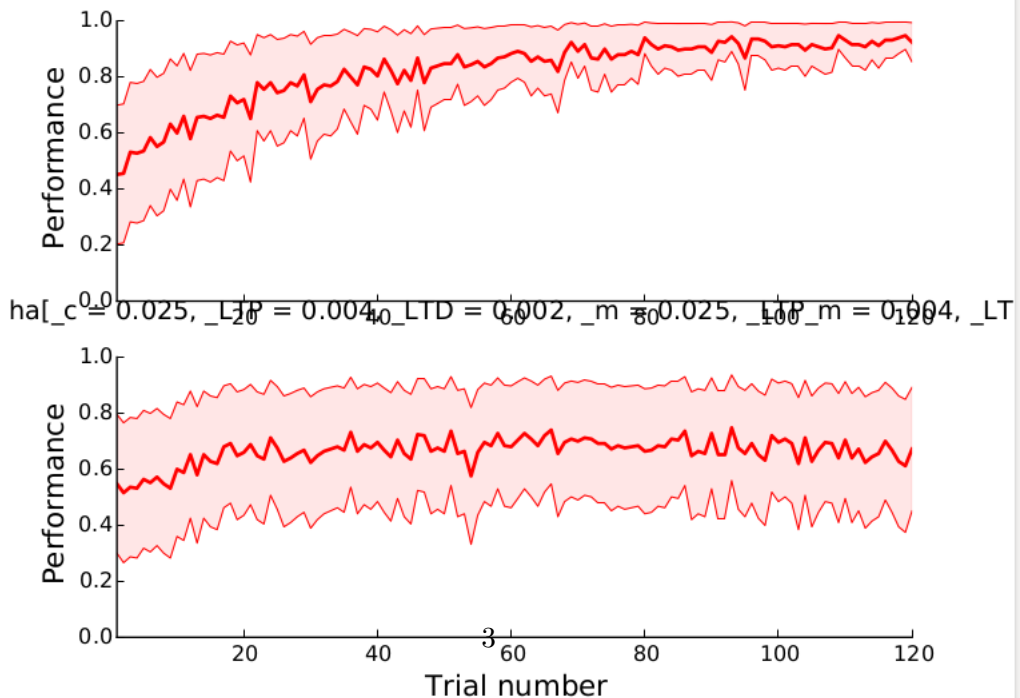
### 2.1.1 Cognitive learning, 250 sessions, 120 trials per session



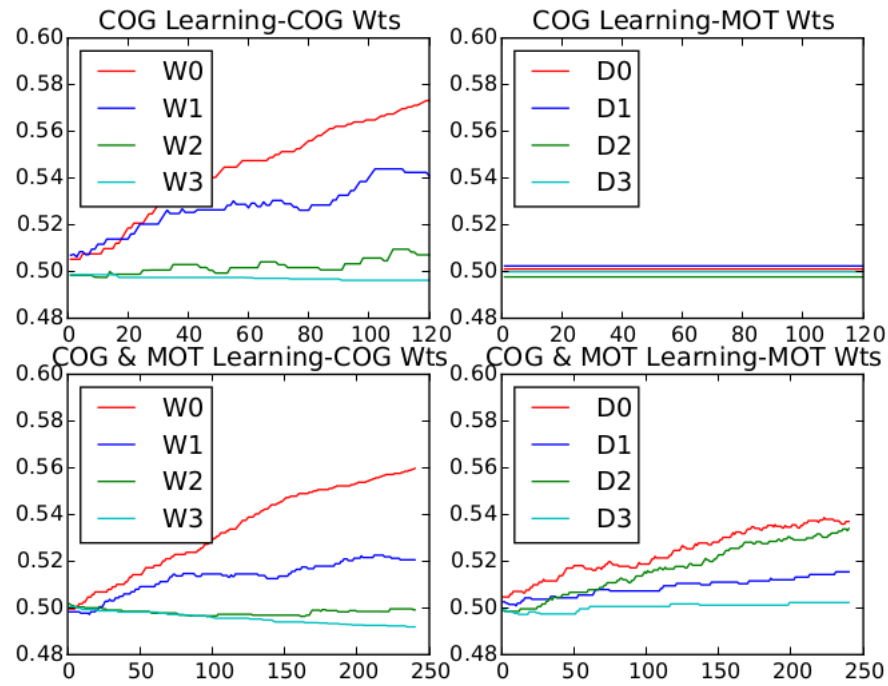
### 2.1.2 Cognitive & Motor learning, 250 sessions, 120 trials per session, standard learning rates



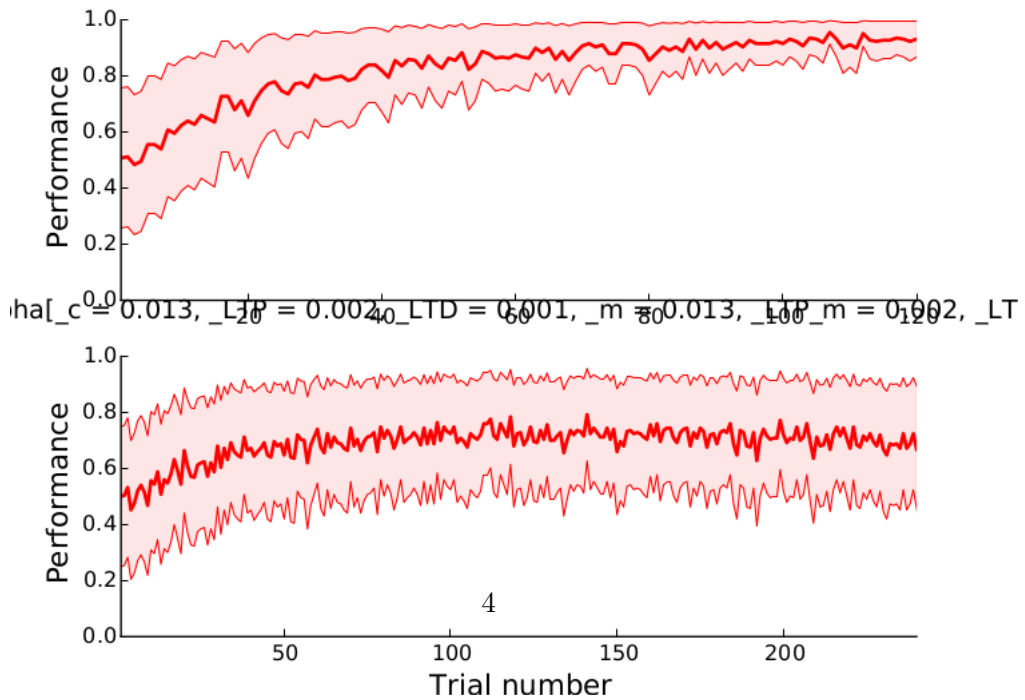
$\alpha[_c = 0.025, \_LTP = 0.004, \_LTD = 0.002]$



### 2.1.3 Cognitive & Motor learning, 250 sessions, 240 trials per session, all learning rates cut by half when there is motor learning



$\alpha[_c = 0.025, \_LTP = 0.004, \_LTD = 0.002]$

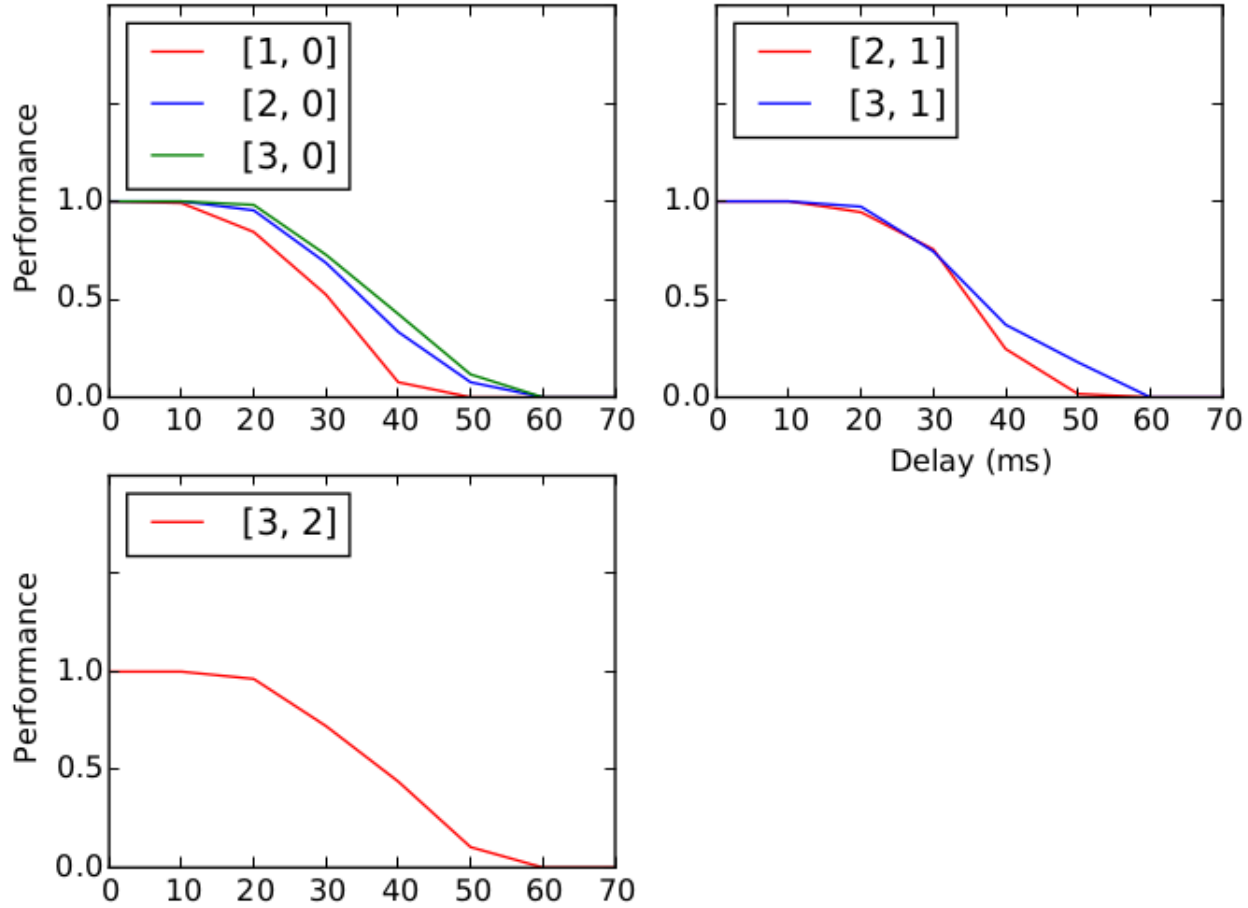


## 2.2 Rewarding stimulus delayed

In this scenario, of every possible combination of 2 cues, less rewarding cue is presented first. After a certain amount of delay, cue more rewarding than the first, is presented. This scenario is tested on a model which is already learnt and which has a performance nearly equal to 1 after learning. Each cue pair is presented 120 times, and mean performance is measured based on the cue selected.

It is observed that the delay after which the best rewarding cue is not chosen varies with the difference of the reward probabilities of the two cues presented.

### 2.2.1 Performance 120 trials/stimulus

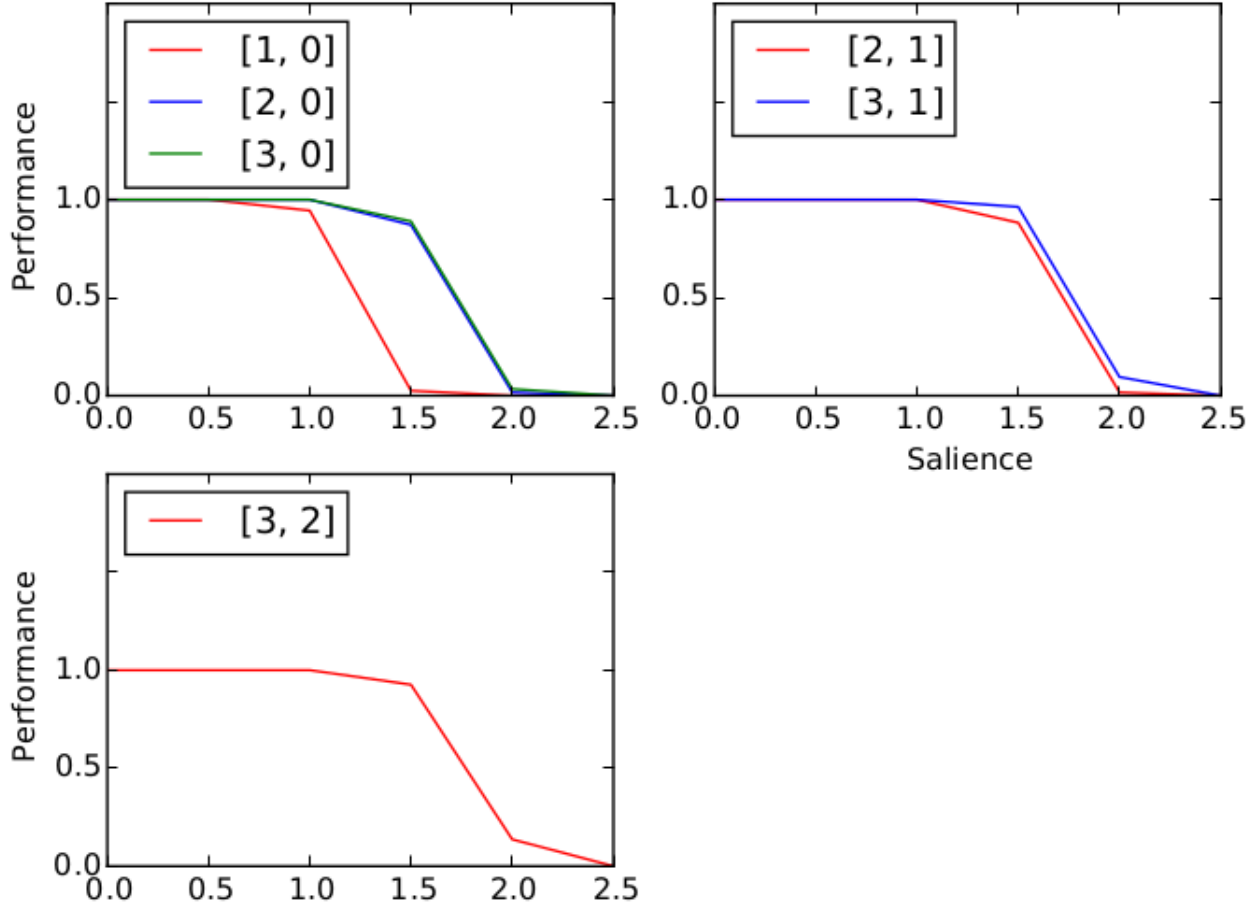


### 2.3 Less rewarding stimulus with higher salience

In this scenario, we consider that a less rewarding (after having learnt) stimulus may be visually more appealing because of its salient features. Both the cues are presented simultaneously, but the less rewarding cue is presented with more salience. In the model, this is achieved by increasing the external input current to the cognitive input of the network corresponding to the less rewarding cue.

Similar to section 2.2, it is observed that the salience level after which the less rewarding cue is chosen over the best, varies with the difference of the reward probabilities of the two cues presented.

#### 2.3.1 Performance 120 trials/stimulus

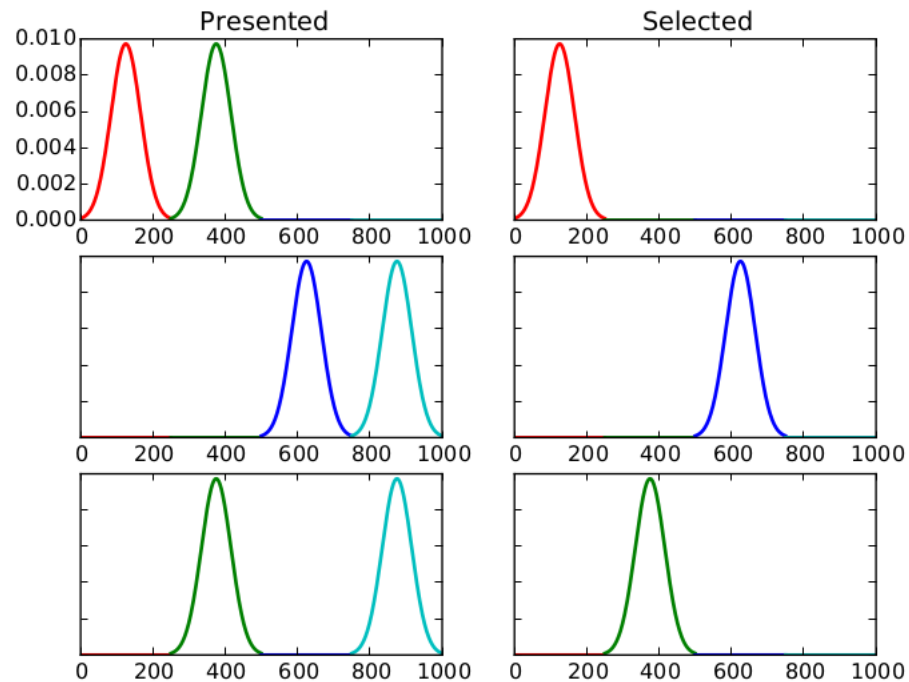


### 3 Week : 10-08-2015

#### 3.1 Scaling up the model

So far, the model represented each cue and each direction as a single unit which is a very simplified way of representation. To scale up the model, the idea is that a large number of *neurons* encode each cue input and direction input. This refers to expanding the structures in the model such as the cognitive channels and motor channels in both cortex (CTX) as well as striatum (STR). Eventually the size of the associative channels in both CTX and STR becomes very large.

The cue input presentation is chosen as a gaussian input over the sub-group of *neurons* representing particular cue. The output is expected to be a similar activation, but only in the sub-group of neurons representing the desired cue output.



#### 3.2 TBD next week

The plan is to identify all the channels and structures in the model which can be scaled from a simplified representation to a more realistic population of neurons