



# Factors affecting reward based decision making : a computational study

Bhargav Teja Nallapu<sup>1,2,3</sup>, Bapi Raju Surampudi<sup>1</sup>, Nicolas P. Rougier<sup>2,3,4</sup>

- 1 - International Institute of Information Technology, Hyderabad
- 2 - INRIA Bordeaux Sud-Ouest Research Center, Talence, France
- 3 - IMN, Université de Bordeaux CNRS UMR 5293, Bordeaux, France
- 4 - LaBRI, Université de Bordeaux CNRS UMR 5800, Bordeaux, France

Contact : bhargav.teja@research.iiit.ac.in



## Abstract

Basal Ganglia (BG) are known to be responsible for action selection, decision making and reward based learning in a changing environment. Using a biologically plausible model, we have been investigating some external and internal factors related to the stimulus representation that might affect the decision making and action selection. We used a computational model of the cerebral structure BG, inspired and replicated from Guthrie et al, 2013 and a two-armed bandit task described in Pasquereau et al. 2007. The task is a probabilistic learning task where stimuli are 4 different shapes associated with different reward probabilities upon selection. At a time, two of the shapes are presented in two distinct positions and the model is expected to make an action to select one of the presented shapes. Upon repeated trials and presented reward after each selection, the model learns the best rewarding cue and is expected to choose the best rewarding cue always thereafter.

One of the questions we attempt to address is to what extent the physical properties of the stimulus like its visual salience, affect the decision to overcome the impact of reward associated to the stimuli. Early results show that there can be an influence of some external and internal factors leading the model to take a bad decision when the worst choice (less rewarding) is presented before the best choice (more rewarding) or the worst choice is more salient than the best one or even if the model learns the reward probabilities associated not to the cue shapes, but to the position where the stimulus is shown.

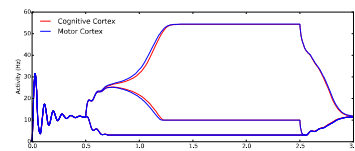
**Model:** A cortico-basal ganglial model with 5 structures. These structures are connected through cognitive and motor decision loops. Each stimulus is represented by one neuron, at every structure.

**Neuronal dynamics:** Each neuron in the model is a rate coded neuron, governed by the equation:

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

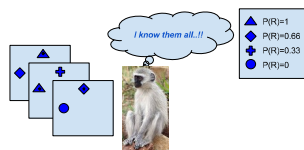
decay time constant of the synaptic input  $\tau$ , negative values of activation,  $m$  and threshold of the neuron  $T$  are set to respective constant values as per Guthrie's model

**Task:** The task is a probabilistic learning task as described in Pasquereau et al., 2007 where four visual cues are associated with different reward probabilities (0.00, 0.33, 0.66 & 1.00). A trial is made of the simultaneous presentation of two random cues with equal salience at two random positions. Some time after the presentation, a switch in the cortex activities is observed, representing the decision taken. After the model has chosen one cue or the other, a reward is given according to the probability associated with the chosen cue. Connections between the cortex and the striatum are then modified using a reinforcement learning rule based on the reward signal. The model is trained over 120 trials such that each combination of cues is presented equal number of times at uniformly sampled positions and the model performance reaches at least 0.9 measuring the ration of optimal choices. The decision switch and the performance are identical to the results when primates are tested with same task.



Cortex activity of all the channels over a single trial. After both the stimuli are presented after 500 ms, a switch in the activity of corresponding channels can be observed, leading the model to take a decision.

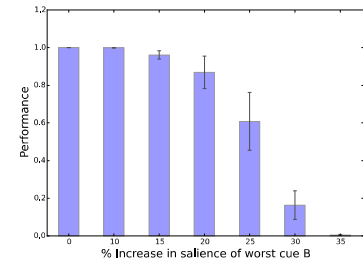
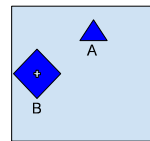
After learning based on the reward received after each trial, the model learns to choose the best rewarding stimulus always. When tested with same task, primates performed similar to the way model predicted.



**Learning** is then disabled and the model is tested using always the same pair of cues A ( P(R)=1 ) and B ( P(R)=0.33 ) in the presence of external factors. We study how, despite reward based learning, visual salience of the stimuli and the temporal difference between stimulus presentations affect the model to take a sub-optimal decision.

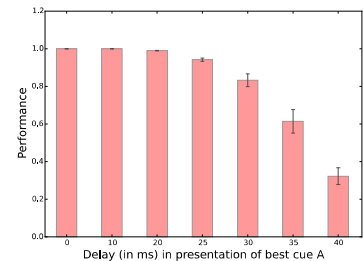
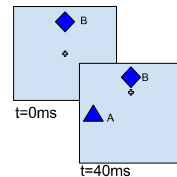
**Factors :** In **scenario 1**, the stimulus known to be lesser rewarding is presented with higher visual salience than that of the higher rewarding stimulus. In **scenario 2**, only the lesser rewarding stimulus is presented at the beginning of trial, and few milliseconds later, the higher rewarding stimulus is presented.

## Scenario 1



**Figure 1 :** Cue 'A' has the best rewarding probability of 1 and cue 'B' is less rewarding, with a probability of 0.33. The model is correct if it selects A when presented with A and B. When stimulus B is presented with higher salience than A (bigger in size, for illustrative purposes), the performance of the model decreases as salience of B increases. It can be observed that when the salience of B is 30% more than that of A, performance of the model decreases to as low as 0.20

## Scenario 2



**Figure 2 :** Cue 'A' has the best rewarding probability of 1 and cue 'B' is less rewarding, with a probability of 0.33. The model is correct if it selects A when presented with A and B. When only stimulus B is presented at first, then after a certain delay stimulus A is presented, the performance of the model decreases as the delay increases.

**Results :** Results furnished above are of the model. The model is already tested to have a consistent performance greater than 0.9 after 120 trials of learning, under conditions of simultaneous presentation of two equally salient stimuli. These early results suggest that such external factors could reduce the performance of the model. Primates have not been tested yet under the scenarios presented here.

**Conclusions :** The decrease in performance explains why primates can never achieve the best performance as the stimuli in reality are under such influencing factors. However, it raises the following questions about the scope and nature of decision making:

- What is an appropriate representation of a stimulus and its salient features (size, color, shape etc.) and their role in overlooking the value of the action while decision making?
- When there is only one stimulus, having known there could be a higher rewarding one
  - what is the time before which the higher rewarding stimulus should be shown, that it could be chosen?
  - how is the decision making process altered when the second stimulus is presented?

**Ongoing work :** In our ongoing work, we have been working on a similar model with a larger population of neurons in each structure. Such a larger population model is not only biologically more plausible, but also provides for more extensive input representations and requires realistic connectivity. Early results show that such a largely populated model is able to learn and perform at around 0.8 accuracy, with certain connectivities assumed.

**Acknowledgements :** The authors would like to acknowledge the funds received from Centre Franco-Indien pour la Promotion de la Recherche Avancée (CEFIPRA) under the project DST-INRIA 2013-02/Basal Ganglia.

## 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 Bezaud, 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.