

Neural mechanisms of human decision-making

Kai A. Krueger (kai.krueger@colorado.edu)
eCortex Inc, Boulder CO; University of Colorado Boulder

Seth A. Herd (seth.herd@gmail.com)
eCortex Inc, Boulder CO

Ananta Nair (ananta.nair@colorado.edu)
eCortex Inc, Boulder CO

Jessica M. Mollick (jessica.mollick@yale.edu)
Yale University, New Haven CT

Randall C. O'Reilly (randy.oreilly@colorado.edu)
University of Colorado Boulder; eCortex Inc, Boulder CO

Abstract

This work presents a neural network model and theory of cognitive decision-making. It attempts to explain the interactions of cortical and subcortical mechanisms and how these may lead to some of the behavioral properties of flexible complex decision-making. Specifically, we model the interactions among several layers of cortex, basal ganglia and the dopamine system. Key to our theory is that all cognitive decisions result from cortico-striatal-thalamus loops akin to those heavily studied in animal motor action selection. Relevant areas of cortex propose a plan of action using associative mechanisms, driven by reinforcement learning; then other cortical areas use that information from sensory input, contextual information and internal goal states to make a prediction about outcome. That prediction is used by striatum to make a go/nogo decision on that plan. Cortical areas thereby learn in a supervised way from actual observed outcomes, whereas the basal ganglia learns its go/nogo decision based on dopaminergic reinforcement signals. By breaking up complex decisions into sequential, simple go/nogo decisions, the same canonical decision-making circuit, as used in basic action selection, can scale up to flexible complex decisions. Furthermore, we postulate that model-free and model-based decision-making are different modes of the same canonical decision-making circuit.

Keywords: Decision-making; Dopamine; PFC; basal ganglia; model-based/model-free

Introduction

Decision-making is of critical importance as the quality of people's decisions are among the most important determinants of whether our outcomes are good, bad, or disastrous. As such, a great deal of scientific work has been directed at human decision-making. We extend this work by providing a systems level neural network model of basal ganglia (BG), the dopamine system and parts of cortex involved in developing plans and predicting outcomes. Our model and theory are

based at the broad level on the close analogy between the brain regions involved in motor action selection, and those involved in higher cognition. In each case, there are circuits descending from cortex, through the striatum, the globus pallidus, and the thalamus, back to cortex. In the motor cortex, these circuits have been studied in great detail, and have been shown to select among competing representations of potential actions in the cortex. Previous computational models of working memory (Frank, Loughry, & O'Reilly, 2001) have proposed that similar basal ganglia circuits connected to higher areas of prefrontal cortex select which representations are maintained by having those neurons enter a state of continuous firing over short periods. We propose that selecting representations in higher areas of frontal cortex serves as an intermediate step in complex decisions, and as the final output for decisions on plans and strategies. It has long been proposed that such actively maintained representations serve to guide thought and action by providing a top-down biasing effect (e.g., (Miller & Cohen, 2001)). As such, they are suited to serve as plans and subgoals in complex tasks. Here we model a laboratory decision-making task with different payouts depending on goals of the agent, and show how the different neural mechanisms interact to evaluate and select appropriate decisions.

Decision-making as a serial, iterative process

Our theory postulates that complex decisions are broken up into an iterative sequence of canonical go/nogo decisions. This sequentialization allows the decision circuit to easily scale to nearly arbitrarily complex situations and set of actions by simply iterating more often through the same loop (see fig. 1). We thereby postulate that this loop consists of generating a candidate plan of action in cortex. Cortex then further predicts the outcome of that considered plan of action, after which it then evaluates that outcome against the current set of goals to assign the plan of action in a given situation and goal state a specific value (most likely to be in orbital frontal cortex). This information then triggers either a go or a nogo response in basal ganglia. A go decision results in the plan being gated into working memory and executed. In contrast,

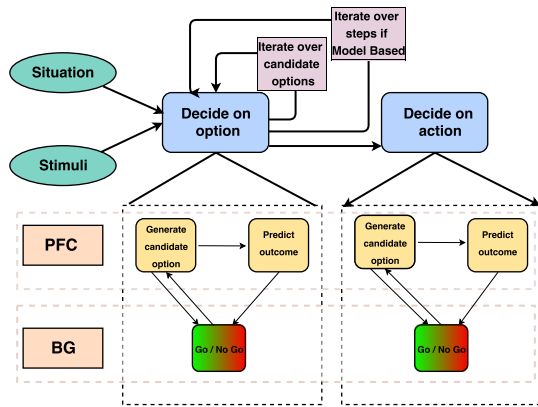


Figure 1: Complex tasks can be broken up into a sequence of simpler go/nogo decisions. Each of these decisions break up into a multi-step pipeline of generating a candidate plan of action, evaluating its outcome and then accepting or rejecting the plan. This same loop iterates through all of plans of action for a given situation, as well as hierarchically loops through plans of actions as the situation evolves.

a nogo decision results in the loop starting over again with cortex presenting a new plan of action for which a go/nogo decision is made.

Model-free / model-based decision-making as opposing ends of the same decision circuit

Recently, the concept of model-free and model-based decision-making has been increasingly prominent in the field of decision-making (Daw, Niv, & Dayan, 2005). It has been suggested that model-based decision-making is associated with cortical learning and model-free with basal ganglia learning based on dopamine. However, increasing evidence suggests a more mixed picture (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Doll, Simon, & Daw, 2012). Behavioural data indicates that the basal ganglia is also involved in model-based decision-making. Furthermore, functional data from motor action selection and other animal data, indicates a close tie between cortex and basal ganglia. The model presented here, instead postulates that model-free and model-based components are working in tandem in the same canonical decision circuit. We postulate that cortex proposes a candidate action plan using a model-free process. Other areas of cortex then predict the outcome of that plan and its respective value given current goals, which by definition is model-based. Finally basal ganglia makes a dopamine driven reinforcement based go/nogo decision. As learning progresses, the order in which cortex considers plans of action increasingly improves such that the first plan considered is often the best available plan. At that point basal ganglia can start selecting the first candidate plan considered before the model-based evaluation has completed, avoiding the time required to do the full model-based evaluation of the outcome. Therefore, despite using the same canonical decision-

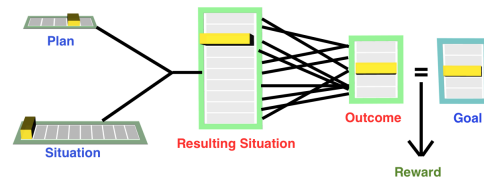


Figure 2: Decision-making task: In each situation, each plan achieves a specific outcome. If the reward outcome matches the goal of the agent, the network receives a positive reward

making circuit, behaviour can shift from more model-based appearance of behaviour to more model-free appearance of behaviour depending on situation and progression of learning.

The task

We simulate decision-making in the model for a relatively complex task intended to capture important aspects of many domains of human decision-making. Three factors must be considered to determine whether the outcome will be rewarding. Those factors are *Situation*, *Plan*, and *Goal*. In each trial, the model is presented with a *Situation* and a *Goal*. It then arrives at a plan to evaluate, and use or reject and move on to another plan. Each Plan, when selected, deterministically leads to one *Resulting Situation*. Each *Resulting Situation* then leads to one *Outcome*. If that *Outcome* matches the current *Goal*, the model is given a reward input; otherwise, no reward (nor punishment) is given for making that selection. Thus, the models task is to avoid (not select) each Plan that will not produce an Outcome matching the current *Goal*. In this simulation, we used 10 *Situations*, 10 *Resulting Situations*, 5 *Plans*, and 4 *Goals* and *Outcomes*. Thus, there were 200 total ($Situation \times Plan \times Goal$) combinations for the model to learn. Thus, we used a larger task space than animals or most humans could easily master purely by memorizing combinations.

A neural network model of cognitive decision-making

We implement and simulate our decision-making theory in a neural network model simulated in *Emergent* (O'Reilly, 2007). Inputs to the model are presented as one-hot encoding layers for *Situations*, and *Goals*, representing one of 10 situations, and one in 4 goals the network is currently pursuing. These layers are connected to the *Plan Selection* layer, which iterates (due to accommodation) in a learned order through the available plans. The chosen *Plan*, together with the *Situations* layer is connected to the *Outcome Prediction* layer, which is trained to predict the *Resulting Situation*. This prediction then further feeds into a *Value Prediction* layer (likely to be OFC in most cases) together with a *Goals* input layer to predict the value of the predicted outcome. Based on the predicted value

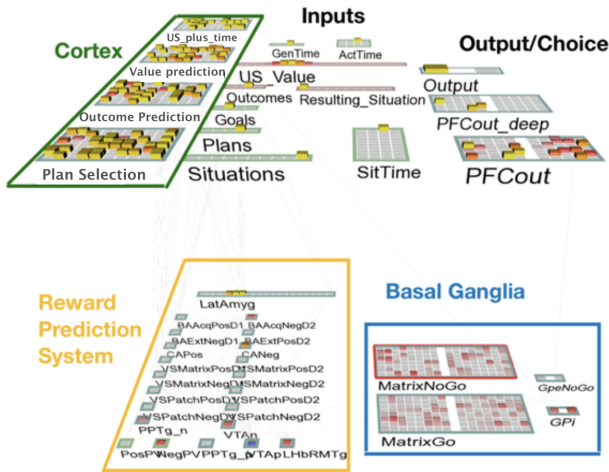


Figure 3: The model consists of 3 main components: Cortex, Basal ganglia and the dopamine based reward prediction system. Cortex generates a plan of action and predicts its likely outcome as well as its value given the models current goal state. Basal ganglia discretizes this information into a go/nogo gating decision. The dopamine system acts as a critic and provides negative reinforcement to indicate a lost opportunity if a plan is chose that doesn't lead to the best outcome.

information, the basal ganglia thalamus loop then chooses to either execute the plan (by the go pathway gating the plan into working memory), or to reject the plan and continuing to consider the next plan of action in the given situation. When a plan is executed, the environment provides a *Resulting Situation*, *Outcome* and the resulting reward signal that the cortical layers use to train their prediction in a supervised way. Furthermore, the overall model incorporates a modified version of the dopamine system that has been developed in the lab over an extended period of time (PVLV) (O'Reilly, Frank, Hazy, & Watz, 2007). This dopamine signal drives learning in the two model-free components of the system, the *Plan Selection* layer, and the basal ganglia go/nogo decision. In each case a positive dopamine signal increases the probability of choosing the plan of action, whereas a negative dopamine signal reduces the probability. Due to the reward prediction error nature of the dopamine signal, it can provide negative reinforcement when a plan is worse than predicted even though the environment in this task only provides positive reward, or no reward.

Results

As shown in figure 4 the model overall learns the task relatively rapidly in the sense that it correctly rejects the action plans that don't lead to reward in a given situation ($92 \pm 1\%$) and accepts the plans that do ($95 \pm 2\%$). The model-based cortical portion of the model learns to predict the resulting situation with $88 \pm 1\%$ accuracy, and the outcome reward with

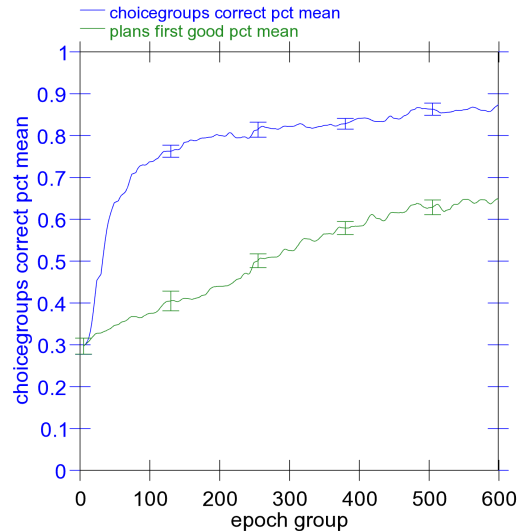


Figure 4: Graph showing performance of the decision-making model on the task. Choicegroups correct depicts the percentage of times the model selects the correct action plan for a given situation and goal. Plans first good depicts the percentage of first plans of action within a choice group that were good plans for the given situation and drive. This part shows the slower model-free learning of considering the correct plan of action initially without needing to evaluate the outcome in a model-based form.

an accuracy of $90 \pm 1\%$. Over time, the *Plan Selection* layer learns to provide a plan that leads to reward as its first candidate plan. Initially it is at chance level $\approx 30\%$ but continues to improve to $65 \pm 2\%$.

Conclusions

Our results show that the full model learns to perform a complex task relatively well, and, over the course of learning, progresses from a model-based to a model-free selection process. While both cortex and basal ganglia are involved in both types of decisions, the cortex contributes substantially less if the model-based cortical prediction components are not performed. These results can explain the findings of relatively more basal ganglia contributions with model-free processing, but activation of both cortex and basal ganglia in both types of processing (Daw et al., 2011; Doll et al., 2012).

This theory does not directly address the functional, descriptive level of many theories based on human behavior. It is intended to provide more specific hypotheses about the mechanisms underlying human decision-making than previous theories have offered. This specificity is possible because we assume (based on well-studied broad anatomical similarities) that the mechanisms of animal action-selection are also those used in human decision-making. We propose that these same basic mechanisms allow humans to make accurate decisions in very complex domains by relying on humans supe-

rior use of abstract representations, and through our ability to assemble many individual micro-decisions into coherent evaluations by use of episodic memory. This theory does not yet fully address how humans are able to use individual carefully-learned evaluative steps into complex chains of logic and prediction in even more complex domains; it remains to future work to determine whether the processes described here are adequate when chained together in sufficient complexity, or whether other, novel mechanisms that are categorically different than animal action-selection are needed.

Acknowledgments

Funding for this project was provided through a sub contract of NIH grant 5R01GM109996-03 (Integrated Cognitive Modelling for Risky Decision-Making) and ONR contract D00014-12-C-0638 (Neural Mechanisms of Human Adaptive Executive Control). We thank all members of the O'Reilly lab for extensive discussions and ideas on this model.

References

- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011, March). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204-1215. doi: 10.1016/j.neuron.2011.02.027
- Daw, N. D., Niv, Y., & Dayan, P. (2005, November). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704-1711.
- Doll, B. B., Simon, D. A., & Daw, N. D. (2012, December). The ubiquity of model-based reinforcement learning. *Current opinion in neurobiology*, *22*(6), 1075-1081. doi: 10.1016/j.conb.2012.08.003
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001, January). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 137-160.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202.
- O'Reilly. (2007). *Emergent Neural Simulation System*. http://grey.colorado.edu/emergent/index.php/Main_Page. (Published: Internet)
- O'Reilly, R. C., Frank, M. J., Hazy, T. E., & Watz, B. (2007). PVLV: The primary value and learned value Pavlovian learning algorithm. *Behavioral Neuroscience*, *121*(1), 31-49.