Integrated Cognitive Architectures For Robust Decision Making

Article · September 2010

CITATION
1

READS
79

4 authors, including:

Christian Lebiere
Carnegie Mellon University
221 PUBLICATIONS **12,433** CITATIONS
SEE PROFILE

Randall C. O'Reilly
University of California, Davis
179 PUBLICATIONS **19,781** CITATIONS
SEE PROFILE

Andrea Stocco
University of Washington Seattle
92 PUBLICATIONS **1,535** CITATIONS
SEE PROFILE

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
September 20, 2010	Final Technical	07/01/2008-11/30/2009

4. TITLE AND SUBTITLE

Integrated Cognitive Architectures for Robust Decision Making

5a. CONTRACT NUMBER

5b. GRANT NUMBER

or FA9550-08-1-0404

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

John Anderson
Christian Lebiere
Randall O'Reilly
Andrea Stocco

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Carnegie Mellon University
4000 Forbes Avenue
Pittsburgh, PA 15213

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR
875 N. Randolph St
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

A = Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

See attached

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			John Anderson
					19b. TELEPHONE NUMBER *(Include area code)* 412-268-2788

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Integrated Cognitive Architectures for Robust Decision Making

FA9550-08-1-0404

Technical Report

John R. Anderson, Christian Lebiere, Randall O'Reilly and Andrea Stocco

Carnegie Mellon University

September 20, 2010

Introduction

The project's goal was to reproduce robust and intelligent decision making capabilities in artificial agents by integrating two successful cognitive architectures, ACT-R (Anderson, 2007) and Leabra (O'Reiily & Munakata, 2000). The rationale was that such an integration effort would yield insights on the general mechanisms that allow rapid decision-making in real-time. Taken separately, ACT-R and Leabra incorporate different views of how decision-making and robust behavior occur. The two architectures have different and complementary strengths and weaknesses and work at different levels of abstractions. Thus, an integration of the two would possibly yield a uniform framework for understanding the computational basis of robust intelligence and decision-making in humans.

Background

Robust decision-making covers the human capability of making choices across different tasks and situations. In this sense, robust decision-making defies the boundaries of traditional machine-learning approaches because it focuses on successful performance with multiple representations and across multiple domains, instead of optimal behavior in a limited domain and with specific data structures.

In our view, two main features characterize robust decision-making. The first feature is *integration*: robust intelligence requires the dynamic integration of different specific cognitive abilities, such as those that allow humans to detect contingencies in the environment and to estimate future rewards. The second is flexibility: robust intelligent behavior requires the capability of dynamically modifying one's own intentions and behaviors in order to match novel changing tasks. Note that integration and flexibility are also general properties of any intelligent and autonomous agent, so that our efforts at unveiling the mechanisms behind robust decision also contributes to the goal of achieving artificial general intelligence.

More specifically, our work has proceeded along two converging lines of research, computational and experimental. Computational research has focused on uncovering and testing possible mechanisms that integrate the decision-making capabilities of two existing general frameworks for modeling human cognition, i.e. ACT-R (Anderson, 2007) and Leabra (O'Reilly & Munakata, 2000). Experimental research has focused on designing tasks that stress the requirements for integration and flexibility in decision-making situation, and collecting relevant behavioral data from human subjects. These two lines of research did not proceed separately, and each progress in one direction spawned novel problems and insights in the other research track.

For clarity purposes, however, we will present the two areas of research in separate sections, occasionally highlighting the intersections between the two.

Part I: Computational Mechanisms for Robust Decision-Making

Integration Between Computational Cognitive Architectures

Robust decision-making is a characteristic of human intelligence. Human cognition can be analyzed at different levels and divided into different fields of research. Among the research that pursues *integrated* theories of human cognition, two approaches have become particularly influencial: ACT-R and Leabra.

ACT-R

ACT-R is a cognitive architecture, i.e., a computational model that aims at providing the basic set of computational operations of the human mind. As in Newell's (1973) original proposal for unified theories of cognition, ACT-R is implemented as a production system, i.e., a mechanism that matches and applies IF-THEN rules. Production systems are a flexible instrument for modeling complex control, and several cognitive architectures have been and are stlll being developed upon them (Just & Varma, 2007; Laird, 2008; Meyer & Kieras, 1997). While production systems are typically symbolic, ACT-R's workings are modulated by a large set of subsymbolic parameters that determine how higher-level symbolic representations are processed are reflect known neural or metabolic processing costs associated to different structures or operations. Thus, ACT-R is not only an integrated theory of human cognition, but also a hybrid theory that reflects the known underlying biological mechanisms.

ACT-R is composed of different modules that provide support for visual perception and attention, motor programming and execution, long-term declarative memory, goal processing, mental imagery, and procedural competence. Perceptual and motor modules are critically important for embodied cognition (the visual and manual modules are illustrated in Figure 1). They enable the system to be "in the world," in contrast to past ACT-R systems where all cognition was in the head. The development of perceptual and motor capabilities has been heavily influenced by the EPIC architecture (Meyer & Kieras, 1997). A critical aspect of learning involves optimizing the perceptual-motor components of the system. A number of applications have now appeared where one or more ACT-R models interact with other agents and complex external systems. These applications depend critically on the ability for external events to interrupt and direct processing.

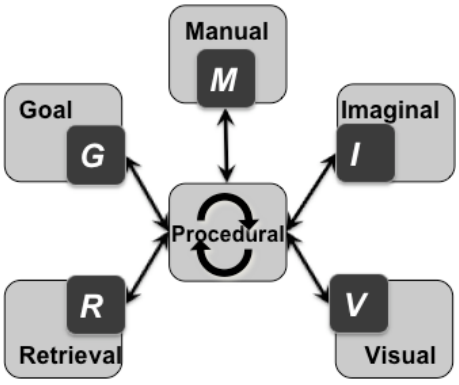

Figure 1: Overview of ACT-R

As can be seen by visiting the ACT-R web site (http://act-r.psy.cmu.edu/), successful models have been developed for a wide range of tasks involving attention, learning, memory, problem solving, decision making, and language processing. Under the pressure of accommodating this

range of tasks the architecture has developed fairly detailed modules. Recent years have seen a major effort to apply detailed modeling approach in ACT-R to the performance of significant real-world tasks. These applications have included driving (Salvucci, 2006), aircraft maneuvering (Byrne & Kirlik, 2005), and simulated agents for computer-generated forces (Best & Lebiere, 2006). We have also continued a long-standing tradition of applying ACT-R models to tutoring systems of academic skills, particularly high school mathematics (Anderson & Gluck, 2001).

In parallel to its application to complex real-word tasks, ACT-R has been significantly extended to predict and incorporate data from the cognitive neurosciences. This has led to an established methodology for predicting the metabolic activity of different brain regions in neuroimaging experiments from the computations performed by different modules (Anderson, 2007; Anderson, Fincham, Qin, & Stocco, 2008). Examples of this work include research on the nature of semantic information retrieval (Danker, Gunn, & Anderson, 2008), problem solving (Anderson, Albert, & Fincham, 2005; Anderson, 2005; Stocco & Anderson, 2008) skill acquisition (Anderson, 2005), and conflict resolution (Fincham & Anderson, 2006; Sohn, Albert, Stenger, Jung, Carter, & Anderson, 2007).

In summary, ACT-R provides an integrated cognitive architecture that is grounded in the basic results of cognitive psychology, and can be dynamically scaled up to model complex tasks, as well as scaled down to examine neurocognitive findings.

Leabra

Leabra (O'Reilly & Munakata, 2000) is the name of a computational framework that has grown out of the original Leabra learning algorirthm for biologically plausible neural network (O'Reilly, 1996). The Leabra algorithm is a learning procedure that integrates two forms of learning: local Hebbian learning and error-driven feedback learning. The integration of the two mechanisms has several advantages over previous learning procedures, and guarantees results.

The core of the large-scale architecture includes three major brain systems: the posterior cortex, specialized for perceptual and semantic processing using slow, integrative learning; the hippocampus, specialized for rapid encoding of novel information using fast, arbitrary learning; and the frontal cortex/basal ganglia, specialized for active (and flexible) maintenance of goals and other context information, which serves to control (bias) processing throughout the system (Figure 2). This latter system also incorporates various neuromodulatory systems (dopamine, norepinephrine, acetylcholine) that are driven by cortical and subcortical areas (e.g., the amygdala, ventral tegmental area (VTA), substantia nigra pars compacta (SNc), locus ceruleus (LC)) involved in emotional and motivational processing. These neuromodulators are important for regulating overall learning and decision-making characteristics of the entire system. Properties of basic neural mechanisms suggest this large-scale specialization of the cognitive architecture.

For example, a single neural network cannot both learn general statistical regularities about the environment and quickly learn arbitrary new information (e.g., new facts, people's names, etc.; McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Rudy, 2001; O'Reilly & Norman, 2002). Specifically, rapid learning of arbitrary new information requires sparse, pattern-separated representations and a fast learning rate, whereas statistical learning requires a slow learning rate and overlapping distributed representations. These properties correspond nicely with known biological properties of the hippocampus and neocortex, respectively. Many empirical studies by ourselves and other researchers, specifically motivated by our computational modeling work, have tested and confirmed these and other more detailed properties. A similar kind of reasoning has been applied to understanding the specialized properties of the frontal cortex (particularly focused on the prefrontal cortex) relative to the posterior neocortex and hippocampal systems. The tradeoff in this case involves specializations required for maintaining information in an active state (i.e., maintained neural firing, supported by the frontal cortex) relative to those required for performing semantic associations and other forms of inferential reasoning (supported by the posterior cortex). The prefrontal cortex system also requires an adaptive gating mechanism

(Braver & Cohen, 2000; O'Reilly & Frank, 2006), to be able to rapidly update some (new) information, such as a new subgoal, while simultaneously maintaining other information that remains relevant, such as the super-ordinate goal. The basal ganglia have appropriate neural properties to provide this function (Frank, Loughry, & O'Reilly, 2001).

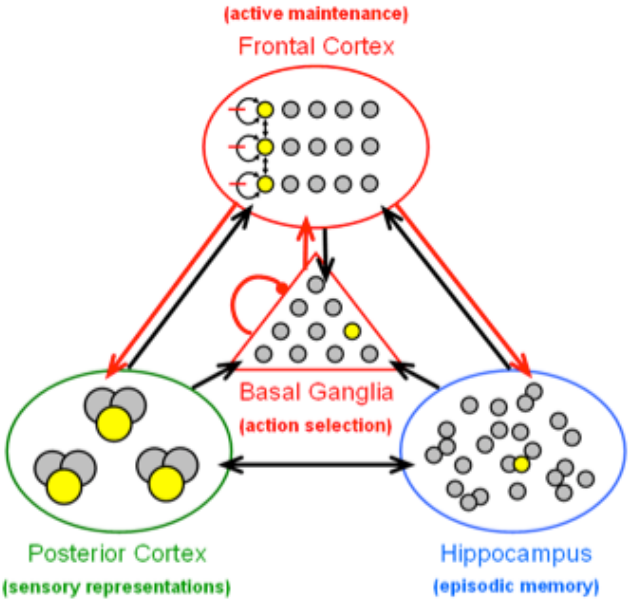

Figure 2: Overview of the Leabra architecture

At the lower level of fundamental neural mechanisms, Leabra integrates into one coherent framework a set of basic neural learning and processing mechanisms (see O'Reilly, 1996; O'Reilly, 1998; O'Reilly & Munakata, 2000; O'Reilly, 2001) that have been otherwise separately investigated in the neural modeling community. Making all these elements work together in a biologically plausible manner is non-trivial, and requires some novel mechanisms, including: a point-neuron activation function that uses simulated ion channels to update a membrane potential with a nonlinear thresholded output to other neurons; bidirectional (i.e., interactive, recurrent) excitatory projections (which are ubiquitous in neocortex) that propagate information throughout the network, integrating information processing on a cycle-by-cycle basis across the different specialized brain areas; inhibitory competition that greatly constrains and speeds the constraint satisfaction process that generates a good representation of the current perceptual inputs; and a synthesis of error-driven, Hebbian, and reinforcement learning, which together produce better overall learning than any of them alone (O'Reilly, 2001; O'Reilly & Munakata, 2000; O'Reilly & Frank, 2006; O'Reilly, Frank, Hazy, & Watz, 2007).

In summary, Leabra provides an integrated and modular architecture for lower-level and biologically plausible models of cognition.

Integration

ACT-R and Leabra share a number of similarities. They are both general-purpose architectures; they are both modular; they share similar views on the mechanisms governing skill acquisition, feedback-driven learning; and they share similar views on the organization of the cortex.

Both ACT-R and Leabra focus on intelligent behavior as the results of integration of information across different specialized modules and different representations. They both share the same overall view of how information processing is distributed and integrated across modules.

The CONDR Model

A first step towards integration was the development of the CONDitional Routing (CONDR) model. CONDR is a model of the basal ganglia that reflects the biology of this circuit and bridges the gap between ACT-R and Leabra.

Transfer of Information in the Brain

Besides ACT-R and Leabra, many ambitious architectures of brain function have been proposed recently (e.g., Houk 2005; Hawkins and Blakesee 2004). These approaches differ widely from each other, but they all have to solve one common problem: the transfer of information among brain regions. The simplest solution consists in hard-wiring the communication between brain regions as direct connections between layers in a network. However, in the human brain, cortico-cortical connections are estimated to make up more than 95% of all the external inputs of a single brain region. Furthermore, about half of this amount is estimated to come from long-distance connections (Braitenberg and Schüz 1991). It is clear, therefore, that some organization needs to be overlaid over this massive set of connections.

The problem of how information is routed between different specialized neural modules deals with the core issues of this proposal, i.e. the issues of integration of *information* and *flexibility* of behavior. It is clear that any computational cognitive architecture that aims at being robust needs to incorporate a mechanism that dynamically allocates and routes signals across modules.

Many different solutions to this problem have been proposed (e.g., Anderson 2007; van der Velde and de Kamps 2006). In this wide space of options, ACT-R and Leabra share a surprisingly common view, as they both propose that the transmission of information along the cortico-cortical pathways is modulated by a subcortical circuit. This circuit comprises important structures such as the basal ganglia and the thalamus. By means of this circuit, organized behavior is imposed upon an otherwise uncoordinated flow of information within the cortex.

Despite sharing this general view, the exact implementation choices taken by ACT-R and Leabra differ in many substantial ways. ACT-R assumes that the basal ganglia functionally correspond to the architecture's procedural module, a structure that provides the long-term repository of production rules and supervises their serial execution. In Leabra, on the other hand, the basal ganglia constitute a gating system that controls the flow of sensory information from the posterior cortex to the short-term memory store located in the prefrontal regions (see Figure 1).

This discrepancy constitutes the biggest obstacle to an integration of the two architectures. Therefore, the first step in our project consisted of devising a model of the basal ganglia that makes the two architectures compatible.

Our solution has been implemented as a connectionist computational model. It provides two additional advantages over other attempts. First, it shows how the subcortical circuit is functionally equivalent to a production system. This equivalence makes an important connection between the anatomy of the brain and a widely studied and adopted computational framework. Second, it provides a natural framework for skill acquisition and habit learning compatible with known biological constraints.

The Role of the Basal Ganglia

Before dealing with computational details, this section will review some evidence in favor of the hypothesis that the basal ganglia play an important role in coordinating the transfer of information between cortical areas. Three converging lines of research support this assumption.

Physiologically, pathologies of the basal ganglia in humans result in an increase in the amount of correlated activity between cortical regions (e.g., Stoffers et al. 2008). This fact can be interpreted by assuming that, under normal conditions, resonance of signals along the cortico-cortical network is limited by the control function of the basal ganglia.

The second line of evidence is represented by studies of human working memory. A number of experiments have shown that the basal ganglia play an important role in gating new information to short-term memory. Neuroimaging data indicating basal ganglia involvement in preparation of working memory updates (McNab & Klingberg 2008). Also, genetic differences in basal ganglia metabolism correlate with individual performance in working memory tests (Zhang et al. 2007). Finally, individual differences in the severity of dopamine depletion in Parkinson's disease also correlate with decline of working memory functions. Control of working memory is the central role of the basal ganglia in Leabra, and corresponds to the role of the ACT-R procedural module in controlling the access to buffers.

A final hint of the basal ganglia role in shaping cortico-cortical connectivity comes from research on learning. It is known that skill acquisition results in a dramatic reorganization of cortical connectivity. Moreover, animal studies have shown that lesions of the basal ganglia result in a profound impairment in skill acquisition (see Packard & Knowlton, 2002, for a review). In animals, it prevents the acquisition of new stimulus-response associations. In humans, it has been proven to disrupt the acquisition of new sensory-motor skills (Cohen & Squire, 1980). Correspondingly, both Leabra and ACT-R, although with different algorithm and implementations, assume that the basal ganglia are required for the acquisition and establishment of new habits.

In summary, experimental research on the basal ganglia has shown their involvement in a number of disparate cognitive functions. This variety of functions can be understood as originating in the circuit's role in coordinating cortical activity. This overarching hypothesis is common to both ACT-R and Leabra.

A Routing Model for the Basal Ganglia

We proposed a model for a brain architecture where the basal ganglia have an overseeing role in directing and shaping cortico-cortical connectivity. This model, named CONDR (CONDitional Routing) is a layered neural network that reflects several aspects of basal ganglia physiology. This network has two key properties. First, it can acquire new stimulus-response associations through practice. Second, its workings can be shown to be substantially similar to a production system. This provides a straightforward mapping between a well-established formalism for artificial intelligence and the biology of the brain.

This architecture works as follows. Let us consider a collection of cortical areas $C = \{c_1 \ldots c_n\}$. For simplicity, let us assume they are all connected to each other. At each moment in time, each region receives signals from $n - 1$ other regions.

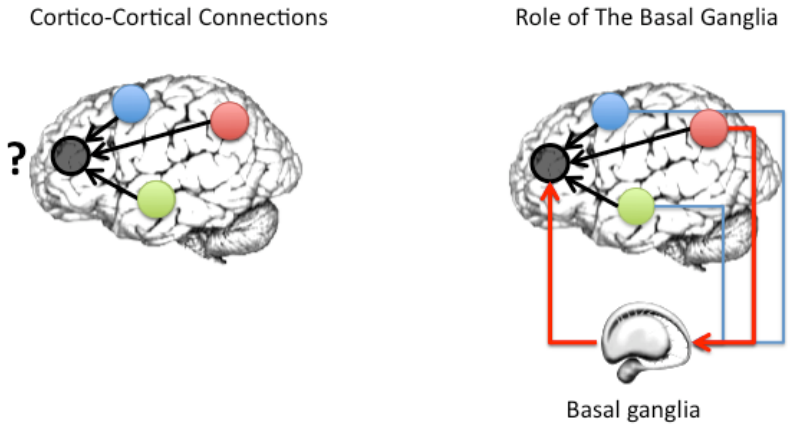

Figure 3: Comunication between massively connected cortical regions (left) can be organized by a cortico-subcortical circuit (right) that encompasses the basal ganglia (bottom right).

Essentially, the basal ganglia alert each region to attend to only a particular subset of "source" regions $S \subseteq C$. This process can be repeated for each region, providing a powerful system for prioritizing and simplifying the exchange of communication (Figure 3). Intuitively, there should be an optimal ratio of $|S|$ to $|C|$. If $|S|$ is too small, the communication between regions is eventually disrupted. If $|S|$ is too large, on the other end, each region receives too many competing signals.

One can consider a very simple model where the pattern held in each region is simply the sum of all the incoming signals from the other regions, $c_i = c_1 + c_2 + \ldots + c_n$.

When the basal ganglia system is compromised, then $S \approx C$, and each region receives almost identical and largely overlapping signals. Therefore, spatial correlation between the regions increases. The temporal correlation increases as well, since updating events in one region are reflected in changes in larger group of connected regions.

Circuitry

In order to explain how the model works, one needs to introduce some biology. The basal ganglia comprise a number of interconnected nuclei. They include the Striatum, the Internal (henceforth, GPi) and External (GPe) Globus Pallidus, the Substantia Nigra (SNr), and the Sub-Thalamic Nucleus (STN). The wiring among these nuclei is usually described in the following terms (Albin, Young, and Penney 1989). The striatum is the entry point of the circuit, receiving afferents from the entire cortex. The nuclei SNr and GPi constitute the system's output. These nuclei project mainly to the thalamus, and control the thalamic projections to the cortex. The Striatum and the SNr/GPI are connected by two pathways, which exert opposite effects. They are known as the direct and indirect pathways. The indirect pathway comprises the GPe and the STN (Figure 4).

A common interpretation, dating back to Albin, Young, and Penney (1989), is that the two pathways simply oppose each other. In particular, the direct pathway conveys excitatory signals to the cortex, while the indirect pathway contrasts this effect through direct inhibition. In the model, we expanded this interpretation as follows. The direct pathway carries a selection of source regions, whose representation has been chosen for transmission. The indirect pathway carries a selection of destination regions for each source region. In practice, the indirect pathway carries a mask that establishes which region each destination should be attending to.

The Striatum

The striatum is the largest nucleus of the circuit. The large majority of its cells are projection neurons (Graveland, Williams, and DiFiglia 1985). These neurons can be divided into two groups: Striatonigral (SN) cells, whose projections form the direct pathway, and Striatopallidal (SP) neurons, whose projections begin the indirect pathway (Figure 4).

In CONDR, the striatum is modeled as a flat structure (see Wickens, 1997) of projection neurons, where SN and SP cells are organized into subdivisions. Each subdivision receives afferents from a single corresponding cortical region. Therefore, the striatal organization mirrors cortical topology. Subdivisions also possess a second-level, internal organization. Within a single subdivision, neurons are grouped into ensembles corresponding to the destination that the source region projects to.

Thus, the model striatal subdivisions reflect cortical topology at two levels. At a macro-level, they mirror the organization of cortex into specific regions. At a lower level, each subdivision also reflects the cortical connectivity of the corresponding cortical region. This organization is compatible with some properties of corticostriatal projection distribution (e.g., Parthasarathy, Schall, and Graybiel 1992).

Physiologically, the activity of projections neurons is controlled by interneurons (IN). Interneurons have an elevated tonic and exert a powerful inhibitory pressure (Tepper and Bolam 2004). Because of the inhibition coming from interneurons, only a small number of ensembles of projection neurons contain active and firing cells. Active neurons in a striatonigral subdivision

signal that the corresponding cortical region is a source region, and its contents have been picked up for routing. Active striatopallidal ensembles within a subdivision, on the other hand, signal destinations where the selected representations should not be transferred.

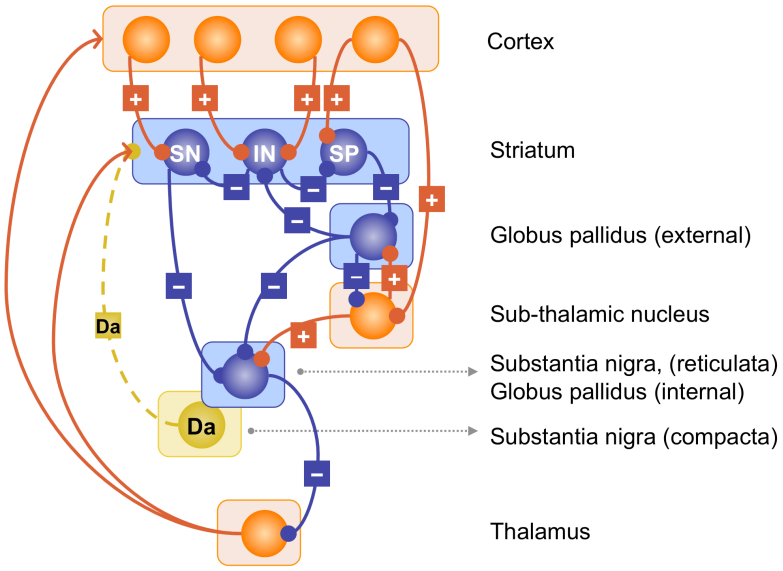

Figure 4 Outline of the basal ganglia circuit and connectivity.

In the model, the dominant inhibitory effect of interneurons was modeling by providing projection neurons with a high threshold θ that is calculated to match the expected incoming signals from the cortex and the inhibitory interneurons:

$$\theta \approx \sum_i w_i E(x_i) \qquad (1)$$

where w_i is weight of the synapses formed with pre-synaptic neuron i, and $E(x_i)$ is the rate-coded expected activation value of i.

Source and destination information are encoded separately in SN and SP neurons, respectively. From there, the two signals travel separately along the direct and the indirect pathway, and eventually combine in the output nuclei SNr/GPi. From the output nuclei, the signals reach the thalamus and, from there, come back to the cortex to enable the proper transfer path. The overall architecture of the model, in a highly simplified rendition, is represented in Figure 5.

Relation to Production Systems

From a purely computational point of view, routing operations can be seen as a neural network analog to production rules in production systems. Production rules are control statements expressed in the form of condition-action clauses (IF-THEN rules). The similarity between the conditional routing model and a production system can be seen if one assumes the following mappings.

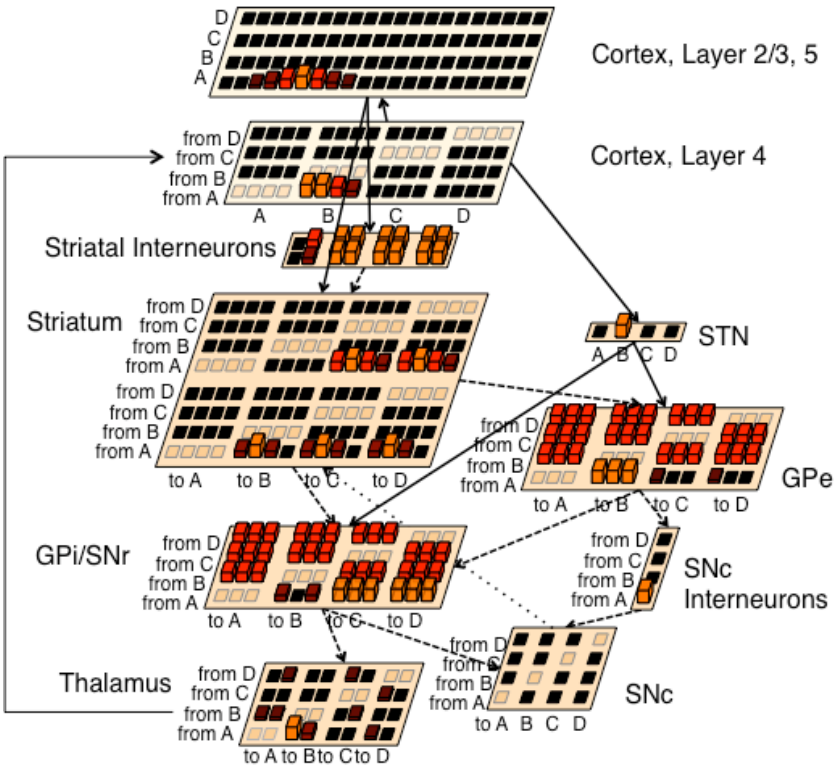

Figure 5: Overview of the CONDR model

A *rule* is embedded in the incoming and outgoing synaptic matrices of a set of striatal interneurons. The *condition* (i.e., left-hand side) part of the rule is represented by incoming synapses to striatal interneurons. The synapses encode the specific cortical representation that will trigger the interneuron to fire. The action (i.e., right-hand side) is encoded in the outgoing synapses to the striatal projection neurons. An action corresponds to the activation of particular ensembles of SN or SP projection neurons in the striatum. Their activation triggers the transmission of information from the source region of the cortex to the destination region.

Part of the flexibility of production systems originates from the use of variables in the production rules. However, variables are not easily dealt with in neural networks. To overcome this problem, a number of procedures have been proposed over the years (e.g., Touretzky & Hinton 1988; Smolensky, 1990; Stewart & Eliasmith, 2008).

In CONDR, the process of biding variables is supported by the special architecture of the model striatum. Consistent with neurophysiology, projection neurons in the striatum are mostly silent, with only a minority of them actually active at any time. In CONDR the active SN and SP neurons correspond to the active combinations of sources and destinations. Ignoring local computations that occur within striatal neurons, the final state of the striatum is the block product $\mathbf{v} \otimes \mathbf{M}$ of the initial vector $\mathbf{v}$ of activations in the source cortical area, and the switchboard matrix of allowed destinations $\mathbf{M}$. The block product can be seen as a special case of tensor product, a powerful mechanism for variable binding in neural networks (Smolensky, 1990). In this case, the variable is the destination cortical region, which is bound to the value v, i.e. the original content of the source region.

In addition to *variables*, production rules can also specify *constants*, which represent default and fixed pieces of information, in CONDR, constants correspond to the transfer of a fixed neural

pattern to a destination region. This particular content is not dependent on any particular source region, and is embedded in the synaptic weights of the circuit. An example of this case will be illustrated in the forthcoming sections on learning.

In summary, the CONDR model provides biological computations that can be akin to the operation of a production system. In particular, the activity of the CONDR model bears similarity with the execution of ACT-R's production rules, where variables are used to bind the contents of a particular destination buffer to the values held in a source buffer. On the other hand, CONDR can be seen as a generalization and expansion of the PBWM model in Leabra. Thus, the model provides an ideal means to integrate the two architectures.

Performance of the CONDR Model

The CONDR model was developed to account for how the basal ganglia can perform general-purpose signal-routing operations, and how their function can provide a flexible way to organize the flow of processing within the cortex. Therefore, the next sections will provide an overview of the model's capabilities and performance. First, we will illustrate how the CONDR model can perform a simple task. The next section will provide evidence of the model's robustness.

CONDR Performance of an Example Task

This section will provide an example of how the model coordinates a series of routing operations to perform a task. The example paradigm is an aural discrimination task that has been used as part of a dual-task experiment by Schumacher et al. (2001) and Hazeltine, Teague, and Ivry (2002). In this task, participants respond to the presentation of a tone. Tones could have three different pitches (220, 880 and 3520 Hz), to which participants had to respond "one", "two", or "three", respectively.

This task requires assembling a number of basic cognitive functions in a novel and arbitrary way, and therefore depends on controlling the flow of information among cortical areas. It is also simple enough that its modeling requires very few assumptions. With some differences in the details, various authors (Schumacher et al., 2001; Hazeltine, Teague, & Ivry, 2002; Anderson, Taagen, & Byrne, 2005) agree that three basic processing steps are taking place: (1) Stimulus classification, during which the stimulus is presented and appropriately encoded; (2) Response selection, during which the appropriate response is selected from the set of possible options; and (3) Response execution, where the chosen response is eventually vocalized.

It is rather uncontroversial that the first and the third step rely on the auditory and motor cortices, respectively (see Anderson, 2007, for an fMRI investigation that confirmed this fact). More uncertain is the localization of response selection. Anderson, Taatgen, and Byrne (2005) proposed an ACT-R model that can successfully reproduce most of the experimental findings. Following ACT-R's mapping of cognitive process onto brain regions (see Anderson et al., 2008), the model implies that response selection recruits the left lateral inferior prefrontal cortex. This interpretation is consistent with the established role of this region in selecting among competing responses in word generation and pair-associate tasks (Danker, Anderson, & Gunn, 2008; Sohn, Goode, Stenger, Carter, & Anderson, 2003; Thompson-Schill, D'Esposito, & Kan, 1999). The specific involvement if this region has been confirmed by an fMRI investigation of this task reported in Anderson (2007, Figure 4.15c).

A simple cortico-basal ganglia circuit was generated to simulate the task. The circuit was simplified to contain only the three cortical regions required by the task. Correspondingly, the striatum only contained three main subdivisions. It was further assumed that each region was connected to the other two. In the model, response selection was simulated as a two-phase step, where the cortical region first attends to the encoded tone from the aural region, and then uses it as a cue to select the appropriate response. To simulate the selection process, the model prefrontal region was connected to a data structure (perhaps corresponding to the hippocampus) that could hold the long-term representations of the three possible responses. The prefrontal

region sends its internal representations to this structure, and receives back the response pattern that is associated with the best-matching input representations. Each cortical region contained 100 artificial neurons.

Figure 6 illustrates how the model performs such task. The figure reads top to bottom, left to right. The four panels on the left-hand side represent the activation of the cortical units, divided into areas, at the three stages of task execution. Note that the two middle panels represent the two phases of response selection. Two routing operations are required to perform this task: they are represented in the two right panels. The two routing operations are required to connect the three task phases. Their implementation follows the model rules described in Anderson, Taatgen, and Byrne (2005). These rules reflect an initial level of task exposure, before participants' performance has been optimized by practice. The model was trained to perform these two routing operations with a Contrastive Hebbian Learning (CHL) procedure.

In Figure 6, panel (a) in the top-left corner corresponds to the state of the cortex when the auditory signal is first encoded. The first routing operation is applied at this stage, and consists in directing the transfer of the tone representation to the prefrontal region. The top right panel represents this routing operation. This panel shows that state of activation of thalamic subdivisions, organized as a source-destination matrix. The active cells are located in the subdivision that projects to the second ("prefrontal") region from the source region ("aural").

Activation of these thalamic terminals determines the transition to the second step, which is represented in panel (b). When the prefrontal region has received the auditory cue, it responds by selecting a pattern corresponding to the response associated to the tone. This phase is represented in panel (c). Note that the model assumes that this operation occurs within the cortex, and the basal ganglia are not involved. The second routing operation (illustrated in the bottom right panel) is triggered at this point, and routes the retrieved response to the vocal region, where it can be executed as a vocal program. This corresponds to the final stage, illustrated in panel (d).

General Performance

Having the model reproducing a particular task does not provide sufficient information on its generality as an information routing device. A series of simulations were therefore carried out to investigate the model's performance. During the simulation, three factors were varied parametrically. Two factors were chosen that affect the model's configuration. They were the numbers of cortical regions (3, 6, 9, 12 or 15, regions), and the size of each cortical region (containing 50, 100, 150, 200, 250 model neurons). The third factor was the number of operations learned before being tested (5, 10, 15, 20, or 25 operations). This factor was chosen to examine the effect of interference between different possible course of actions.

While the size of cortical regions was varied parametrically, the size of each striatal subdivision was kept constant across simulations. In particular, each striatal subdivision contained 20 striatonigral and 20 striatopallidal neurons. Also, the number of neurons in each number thalamic and SNr/GPi subdivision was kept equal to 10. These values were kept constant, so that each different cortical region's size also corresponded to a different ratio of cortex-to-striatum size.

For each training test, a specified number of operations were generated randomly. Each operation was to be performed in response to a different, randomly generated pattern of cortical representations, and the model was trained on each of them. Training was performed by means of a modified version of Contrastive Hebbian Learning. During testing, one of the operations was then selected at random, and its cortical pattern presented to the model. The pattern was propagated through the circuit, and the state of the thalamic subdivisions compared against the desired response as in the previous set of simulations.

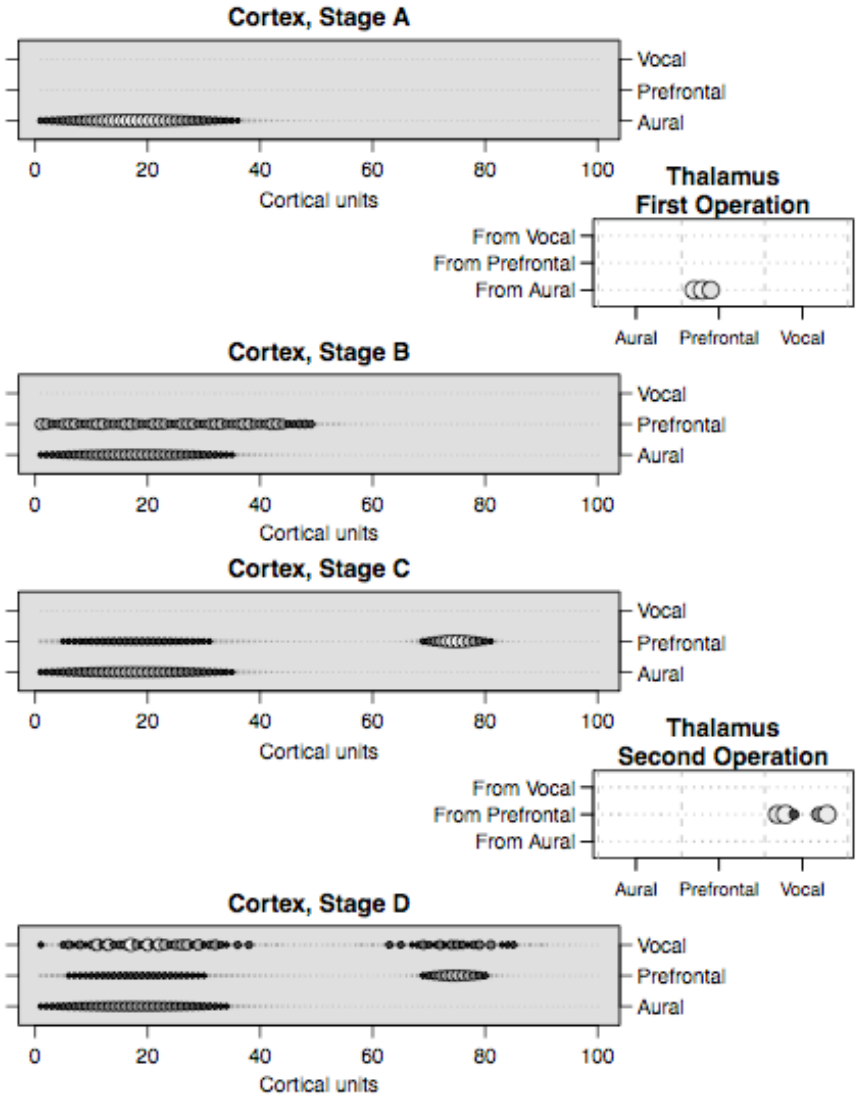

Figure 6: The Aural-Vocal tasks as performed by the CONDR model

The model was tested 100 times for each level of the three factors (number of regions, cortical size, and operation complexity). Trial performance was assessed by comparing the state of the thalamic subdivisions against the desired response. A trial counted as incorrect whenever (a) there were active cells that did not correspond to a proper source/destination binding; or (b) the desired cells were not active.

The number of incorrect trials was counted for each combination of factor levels. Each factor and each two-factor interaction was then analyzed independently, using a fixed-effects statistical model. The size of cortical regions did not have any significant effect on the model's performance [$F(4, 120) = 0.23$, $p = 0.91$], and did not interact with the other factors [$F(16, 100) < 0.55$, $p > 0.92$]. On the other hand, the number of cortical regions [$F(4, 120) = 23.14$, $p < 0.0001$], of routing operations [$F(4, 120) = 2.65$, $p = 0.03$] and their interaction [$F(16, 100) = 6.38$, $p < 0.0001$] were all significant.

The left panel of Figure 7 illustrates the percentage of errors for each combination of number of regions and operations, collapsed across different sizes of cortical regions. It can be seen that the

probability of making an error increased with the number of possible operations, and decreased as the number of regions increased. This increase in errors can be due to the fact that, as the number of regions decreases, the routing operations' patterns become increasingly similar. Under such circumstances, undesired source-destination bindings, which were supposed to be the response of a different operation, might show up in addition to those of the executed operations.

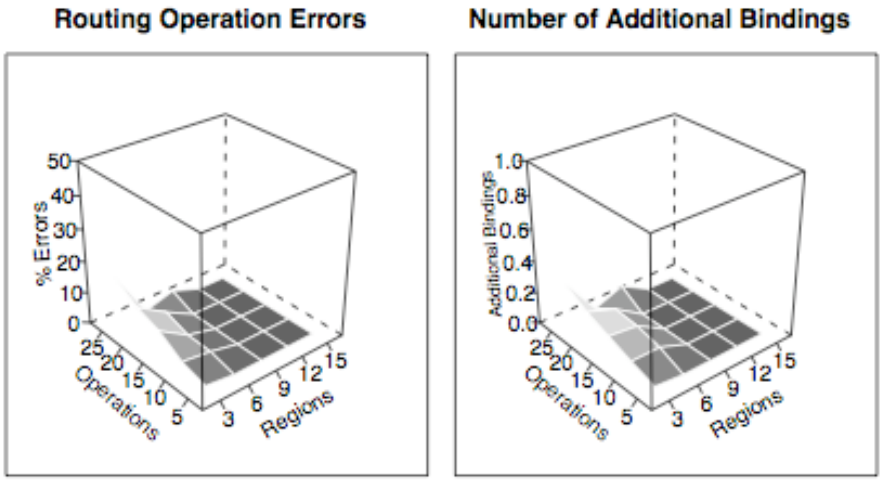

Figure 7: Performance of the CONDR model

To examine this possibility, we analyzed a different measure of model's performance. This measure was the number of additional bindings. An additional binding was defined as a thalamic subdivision that contains active neurons, but does not belong to the desired source-destination bindings. The analysis confirmed our prediction. The number of additional bindings was not affected by the cortical size or its interaction. However, like the percentage of correct trials, this measure decreased with number of regions [$F(4, 120) = 11.52$, $p < 0.0001$], increased with the number of operations [$F(4, 120) = 2.27$, $p = 0.06$], and was affected by their interaction [$F(16, 100) = 2.17$, $p = 0.01$; see Figure 7, right).

Summary

This section has presented an overview of the model's capabilities. The model's capabilities were tested in two different ways. First, it was shown how the model performs a simple stimulus-response task. Second, it was shown how robust the model's performance is when a number of changes are made to its configuration (e.g., increasing size or increasing number of cortical regions), and the number of available responses is progressively increased. Overall, the model's robustness in face of large changes in its structure confirms the efficiency of the basal ganglia architecture for routing information.

Habit Learning and Dopamine in the CONDR Model

One of the most important functions of the basal ganglia is to enable skill acquisition. This capability is particularly crucial in the context of integrating different architectures, as learning enables the incremental acquisition of knowledge and procedures that underpins dynamic and robust decision-making.

In fact, both Leabra and ACT-R incorporate mechanisms for acquiring new skills and procedural knowledge. In particular, both architectures include special mechanism for dealing with feedback-

related and reward-related learning. One crucial difference is that ACT-R, but not Leabra, possesses a mechanism for acquiring new skills with practice, even in the absence of feedback. One of the goals of the CONDR model was to overcome this gap by providing a biological account of practice-related learning. This account could be used to explain ACT-R's properties at a neural level, and eventually integrated in Leabra.

The learning capabilities of the basal ganglia are crucially modulated by dopamine. Dopamine is a neurotransmitter that affects neural plasticity, promoting long-term potentiation and long-term depression among neurons. Some properties of the dopamine signal in the basal ganglia have been successfully modeled as the error term in Sutton's (1988) Temporal Difference algorithm (Schultz 2002). That is, the dopamine signal reflects the error between two subsequence predictions of a specific state's value. Many models of dopamine function have been proposed (see Joel et al. 2002 for a review). Both ACT-R and Leabra provide reinforcement-learning mechanism for reward-related learning (O'Reilly & Frank, 2006; Fu & Anderson, 2004). The difference between these two mechanisms has been the subject of an experimental investigation described in Part II of this report.

More importantly, the connection between dopamine and procedural learning has been less investigated. It is easy to imagine that dopamine also underlies procedural learning. This is particularly important because the exact mechanisms by which new skills and habits can be acquired have seldom been modeled (See Ashby, Ennis, & Spiering, 2007, for a notable exception).

Habit Learning and Production Systems

All the learning in CONDR is due to changes in the strength of synapses between neurons. Computationally, these changes follow simple Hebbian rules. Hebbian algorithms are regarded as a plausible approximation to the biological dynamics of synaptic long-term potentiation and long-term depression (Brown, Kairiss, & Keenan, 1990). In Hebbian learning, changes in synaptic weights (indicated as $\Delta w_{i,j}$) are proportional to the product of pre- and post-synaptic activations. Many variations of this principle have been proposed, differing in mathematical properties such as long-term stability and convergence (Dayan & Abbott, 2001; Gerstner & Kistler, 2002). In our model, the Hebbian rule was implemented as follows:

$$\Delta w_{i,j} = r\,(x_i - \langle x_i \rangle)\,(x_j - \langle x_j \rangle) \qquad (1)$$

where r is the learning rate, and $\langle x_i \rangle$ denotes neuron i's baseline activity. This rule states that the synapses between two neurons are strengthened whenever their firing rates conjointly exceed or fall below their baseline activation. A negative value of r was used for inhibitory projections. This turns the rule into an anti-Hebbian algorithm, which maintains the correct direction of long-term potentiation (LTP) in inhibitory synapses.

Striatal interneurons are dealt with in a special way. These neurons are characterized by a high baseline activity ($\langle x \rangle$), and their firing rate decreases when significant cortical patterns are detected (see Appendix A for the exact mathematical implementation). To account for this asymmetry, the opposite term $\langle x \rangle - x$ (instead of $x - \langle x \rangle$) was used whenever it referred to striatal interneurons.

Specialization in the Example Task

Within our model, this simple form of Hebbian learning is sufficient by itself to capture certain features of skill acquisition, namely, specialization and automaticity. Figure 8 illustrates the changes in the response of SN projection neurons after a different amount of repetitions of the task. The SN neurons belong to the striatal subdivision that receives projections from the prefrontal region (source), and transmits information to the vocal region (destination). In the both panels, time flows horizontally. The left panel details the activation of SN neurons at different

levels of practice, corresponding to 0, 3, 6, or 9 repetitions of the very same trial. Their activation reflects both the excitatory input from the cortex and the decreased inhibition from interneurons. The right panel reflects the contribution of striatal interneurons only, without the cortical component. The figure shows that, with practice, the pattern that is embedded in the synapses between interneurons and projection neurons comes to resemble the incoming cortical input.

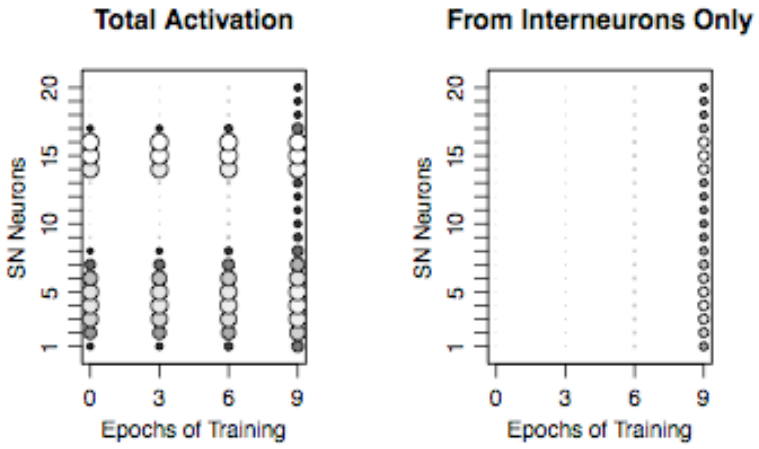

Figure 8: Development of specialized representations in the CONDR model

These changes provide a preliminary basis for the development of automaticity. As long as the same information is available in the striatal interneurons' projections, the original cortical representation is not needed, and the same pattern can be used without the need for cortical processing. This fact is consistent with the drop in cortical activation that can be experimentally observed with practice (Chein & Schneider, 2005; Hill & Schneider, 2006; Raichle et al., 1994; Qin et al., 2003).

Dopamine and Skill Learning

Simple associative learning mechanisms cannot go very far. The modulation of dopamine in the striatum, however, can strategically direct Hebbian learning, significantly increasing the model's learning capabilities. One important way in which practice can improve performance is by eliminating intermediate processing steps that require cognitive control. An example of such processing step occurs between stages (b) and (c) in the example task (see Figure 5). In this step, the prefrontal region uses the auditory stimulus as a cue to retrieve an associated response from long-term memory. With practice, this extra step can be replaced by a specialized routing operation that binds the initial auditory stimuli with their associated responses.

Computationally, the idea of producing novel knowledge by creating direct stimulus-response mappings and skipping intermediate steps has been exploited in a number of production system learning algorithms. These algorithms include powerful techniques like chunking (Laird, Rosenbloom, & Newell, 1986) and production compilation (Taatgen & Lee, 2003). All these techniques have a long record of successes in modeling human learning. Furthermore, they can be seen as examples of skill learning, which is one of the memory functions of the basal ganglia (Packard & Knowlton, 2002). Thus, by providing a mechanism that allows practice-related acquisition of new skill the CONDR model can further reduce the gap between ACT-R and Leabra.

In CONDR this kind of learning takes place when the basal ganglia harvest in one cycle a representation from the very same cortical source (in the example, the lateral inferior prefrontal region) that was the target destination in the previous cycle. The fact that a region that has been recently used as a destination now figures among the source of routed representations is an important cue that this region has been used for some intermediate processing.

Learning to skip steps exceeds the capabilities of the simple Hebbian dynamics presented in the previous section. It can be accomplished, however, by strategically guiding the Hebbian rules. In the model, this is accomplished by the intervention of dopamine, a neurotransmitter that plays a crucial role in changing synaptic plasticity in striatum (Calabresi et al., 2000; Wickens, Begg, & Arbuthnott, 1996). Biologically, the striatum receives dopamine from two major pathways, the mesolimbic pathway from the ventral tegmental area (VTA), and the nigrostriatal originating from the SNc.

Much is known about the response of dopamine neurons to unexpected rewards, and how their bursts closely reflect the reward prediction error (Schultz, 1998; 2002). Although the routing model can, in principle, also learn from these reward-related responses, this paper focuses on a different, practice-related form of learning. This type of learning depends on the release of dopamine under specific additional circumstances. In particular, we hypothesized that procedural skill acquisition is mediated by dopamine signal carried by the nigrostriatal dopamine pathway originating in the SNc. This pathway is essential for habit formation (Faure, Haberland, Conde, & El Massioui, 2005), and has been previously included in other models of basal ganglia (e. g., Ashby, Ennis, & Spiering, 2007).

The SNc receives direct projections from the striatum, as well as indirect projection through the external pallidus and the SNr (Haber, 2003; Haber, Fudge, & McFarland, 2000; Hajos & Greenfield, 1994). Thus, its activity can be modulated by the other nuclei of the circuit. Among these afferents, there is evidence that the influence of direct striatonigral projections is rather weak. For instance, the spontaneous activity of dopamine neurons does not change when the striatonigral connections are removed (Hajos & Greenfield, 1994). Both the projections from the GPe (Hattori, Fibiger, & McGeer, 1975; Smith & Bolam, 1990) and the SNr (Tepper, Martin, & Anderson, 1995), on the other hand, have significant effects on the output of dopamine neurons. The model dopamine system, whose architecture is shown in Figure 9, is also controlled by the SNr/GPi and GPe projections.

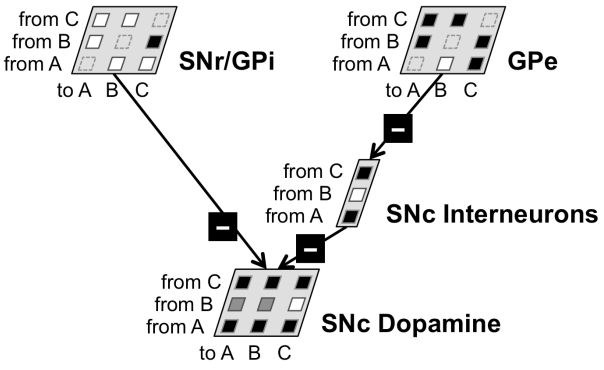

Figure 9: The dopamine circuit in the CONDR model

Dopamine projection neurons in the SNc receive direct input from the SNr/GPi as well as from local SNc interneurons (Hajos & Greenfield, 1994). The proposed mechanism additionally

assumes that the SNc interneurons receive projections from the GPe, and that their activation persists for a sufficient time to provide a delayed memory of the destinations that were used in the previous transfers (Figure 3). This delay is crucial because it allows the SNc dopamine neurons to directly compare the sources of the representations that are being transferred (SNr/GPi) with the destinations of the previous transfers (the delayed signal from GPe). Note that the model does not assume that the GPe signal itself delayed with respect to the SNr; it simply assumes that the GPe signal can be temporarily maintained for comparison with the subsequent patterns of activity from SNr. The discussion section examines other possible mechanisms that produce the same effect. Because of this delay, when the current sources figure among the previous destinations, the sum of inputs from the SNr and GPe triggers an increase in dopamine output to the striatum. This case is visually represented in Figure 9.

There are different ways of modeling the effects of dopamine. Our model adopts a solution proposed by Ashby, Ennis, and Spiering (2007), who captured the role of dopamine by adding a third term to the Hebbian rule for striatal synapses. This third term reflects the activity of dopamine neurons, and expresses the biological fact that learning in the striatum is due to the interaction of pre- and post-synaptic neurons with dopamine (Ashby, Ennis, & Spiering, 2007; Miller, Sanghera, & German, 1981). Within our simple Hebbian framework, this three-way interaction can be easily reproduced by including the activation x_d of dopamine neurons to equation (1):

$$\Delta w_{i,j} = r\,(x_i - \langle x_i \rangle)\,(x_j - \langle x_j \rangle)\,(x_d - \langle x_d \rangle) \qquad (2)$$

If we indicate $(x_d - \langle x_d \rangle)$ with the symbol d, we can re-write equation (2) as:

$$\Delta w_{i,j} = d\,r\,(x_i - \langle x_i \rangle)\,(x_j - \langle x_j \rangle) \qquad (3)$$

Equation (3) makes it apparent that increases and decreases of dopamine modulate synaptic plasticity by increasing or decreasing the learning rate. In fact, when dopamine falls below baseline (i.e., $d < 0$), the direction of learning can even be inverted. Although this equation does not capture all the subtleties of learning in the basal ganglia, it has the advantages of being simple and free of additional assumptions. Therefore, dopamine effects on learning were modeled by increasing or decreasing the learning rate term dr.

In addition to modulating the learning rate, dopamine directly affects the activity of striatal cells. In particular, it excites SN neurons and inhibits SP cells (Bolam et al., 2000; Nicola, Surmeier, & Malenka, 2000; see Figure 2). These differential effects permit a fine modulation of the direct and indirect pathways, which has often been included in basal ganglia models (Frank, Loughry, & O'Reilly, 2001; O'Reilly & Frank, 2006). Less frequently modeled, but equally important, are the opposing effects of dopamine on GABAergic and cholinergic interneurons (Tepper & Bolam, 2004). Our model contains one single type of interneuron that captures properties of both. Since cholinergic interneurons also control the fast-spiking GABAergic interneurons (Figure 2), their reaction to dopamine was taken as the dominant one. Therefore, dopamine inhibits the model interneurons. Excitatory and inhibitory effects of dopamine neurons were modeled by simply using excitatory or inhibitory projections from SNc dopamine neurons to striatal cells.

Skill acquisition depends on the dynamics between all these effects. Long-term potentiation in the corticostriatal projections increases the probability of SN neurons firing, and of interneurons to deactivate in presence of a similar pattern of cortical activity. After repeated exposures, the synapses between interneurons and projection neurons encoded the transferred representation (see Figure 7); therefore, this representation can be imposed to SN neurons even in absence of

the original cortical input from the corresponding region.

Skill Learning in the Example Task

The simple aural-vocal task described in the previous section is useful for demonstrating these effects of learning. In the simple model outlined above, the correct response to a tone had to be selected from long-term memory (see Figure 6, stages (b) and (c)). This intermediate step can be omitted with practice. The lateral prefrontal cortex figures as the destination of the first routing operation (Figure 5, top-right panel) and as the source of the second (Figure 6, bottom-right panel). Therefore, the redundant step can be detected in the convergent pathways on the SNc neurons. In turn, this triggers dopamine release in the striatum, initiating the learning process described above. Figure 10 illustrates how the model performs the task after the learning step has happened a sufficient number of times to allow the newly learned routing operation to fire. The figure illustrates how the new operation routes an immediate response to the vocal region when presented with the original stimulus. Thus the model transitions from the initial stage to the final stage in Figure 10 without the intermediate stages in Figure 6.

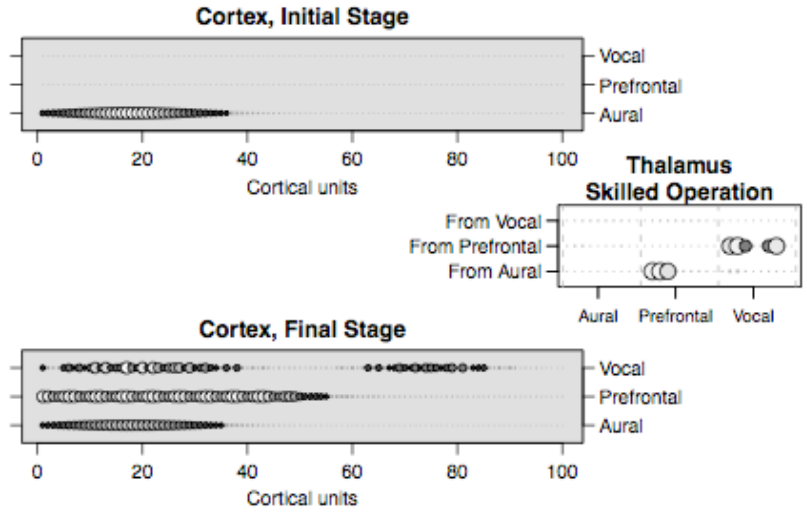

Figure 10: The Aural-Vocal tasks as performed by the CONDR model after learning has occurred

This form of learning relies on the strategic release of dopamine after the execution of the second operation. We have outlined one possible biological mechanism by which this could happen. It should be noted, however, that other mechanisms of dopamine control could obtain similar results. One of these mechanisms is the simple release of dopamine according to unpredicted rewards (Schultz, 1998; 2002). Thus, an increase of the d term after the second operation can also be triggered by the initial reward generated by succeeding in the task.

Summary

This section has illustrated the learning capabilities of the model. In particular, it has shown how the model, with practice, can encode internally certain representations that were originally routed from the cortex. This accounts for the proceduralization and the specialization of responses. It was accomplished by means of simple Hebbian learning, which is a biologically plausible learning rule. When coupled with the effects of dopamine, the Hebbian rule triggers more complex dynamics. Eventually, these dynamics enable the acquisition of new skills that skip intermediate steps in series of information transfers. The elimination of redundant steps during the learning phase accounts for practice speedup and automaticity. Finally, the time course of striatal activity is consistent with important established results in the field of habit learning.

Detecting the appropriate conditions for dopamine release requires a comparison of the current sources and the previous destinations in the SNc. In turn, this requires maintaining a delayed version of the previous destinations in the SNc interneurons. Although this account is partially speculative, many other models have adopted and defended similar mechanisms that compare the current signals from the direct pathway with a delayed signal from the indirect pathway (e.g, Barto, 1995; see Joel, Niv, & Ruppin, 2002, for a review). It should also be noted that similar results could be also accomplished by other mechanism that regulate dopamine release. For instance, reward-related changes in dopamine also follow a pattern similar to Figure 10, with dopamine decreasing as a task becomes more practiced and rewards become predictable (Schultz, 1998; 2002).

Summary: The CONDR Model

This paper has presented a connectionist model of the basal ganglia. The model is based on the idea that this subcortical circuit oversees and controls the transfer of information among cortical brain regions. Because of the large underlying amount of cortico-cortical connections, this circuit plays a fundamental role in organizing the flow of information within the brain by selecting proper source and destination regions. The model can be seen as a neural implementation of a production system, where production rules correspond to routing operations among brain regions. This equivalence is important for two reasons. The first is that it provides a biological substrate for a powerful and well-known computational framework. Second, this equivalence provides a means to understand the neural basis of intelligence and flexible behavior, and bridge the gap between low-level and high-level computational descriptions of the brain.

Instruction Following

One of the hallmarks of robust and intelligent behavior is the capability of directing one's own behavior on the basis of predefined, declarative representations. This capability is useful because declarative knowledge is usually more flexible to manipulate than other types of knowledge, and can be more easily communicated. Humans routinely exhibit this type of intelligent behavior when they are engaged in complex tasks such as planning or problem solving. Perhaps the most striking example of this behavior is following instructions, i.e. the capability of translating abstract representations of behavior into action. This process is akin to interpreting a programming language statement in computer science. Computationally, this process requires some mandatory computational steps that are independent of the implementation of the interpreter itself; in particular, instructions need to be translated into operations and structures that match the underlying hardware.

The process of following instructions has been successfully modeled in ACT-R (see Taatgen, Huss, Dickinson, & Anderson, 2008). In ACT-R, it is possible to represent abstract behaviors as chunks of declarative knowledge, and use special production rules to interpret and instantiate them. The same process, however, is not easily reproduced in a neural network, and therefore this powerful mechanism does not scale down to a Leabra-like architecture. Furthermore, there is no direct evidence that links this process to a clear neural substrate.

In this second part of our computational modeling research, we have provided computational evidence that the basal ganglia are crucially recruited in instruction-interpreting behavior, and that the CONDR model provides an ideal platform for implementing this mechanism within a biological neural network. An experiment validating the model's predictions is discussed in the Experimental section.

The Task

Instructed behavior is seldom investigated in cognitive psychology, and data from the instructional phase of experiments routinely discarded. Thus, we developed a novel task that was used for both testing our models and collecting experimental data from participants. The task consists in

solving a series of arithmetic problems, each of which is combination of three operations, such as "divide x by 3", "multiply y by 2", and "multiply x and y". Each problem required exactly two input numbers (x and y) and always contained one binary and two unary operations. In order to ensure that intermediate and final results were always integer numbers, participants were instructed to use the quotient as the result of a division, and discard the remainder (e.g., 7 / 2 = 3). The three operations were randomly selected from a set of five, each of which was associated to an alphabetical letter *A*, *B…E*. Table 1 illustrates the operations used in the experiment and provides some examples.

Each trial consisted of three consecutive phases: (a) An instruction phase, where the problem was presented; (b) An execution phase, where the two input numbers where presented and calculations were performed; and (c) A response phase, were participants indicated whether a certain number was the solution to the problem or not. The structure of a sample trial is illustrated in Figure 11.

Instructions were presented as a string of letters and variables such as *AExDy*. Instructions were in prefix notation, so that the above problem was interpreted as $A(E(x), D(y))$, which can easily be translated as $(x / 3) \times (y + 1)$ by looking at Table 1.

Table 1: The five operations used in the experiment

Operation	**Meaning**	**Examples**	
$A(x, y)$	$x \times y$	$A(4, 2) = 4 \times 2 = 8$;	$A(2, 3) = 2 \times 3 = 6$
$B(x, y)$	x / y	$B(8, 2) = 8 / 2 = 4$;	$B(6, 3) = 6 / 3 = 2$
$C(x)$	$x \times 2$	$C(4) = 4 \times 2 = 8$;	$C(3) = 3 \times 2 = 6$
$D(x)$	$x + 1$	$D(7) = 7 + 1 = 8$;	$D(3) = 3 + 1 = 4$
$E(x)$	$x / 3$	$E(9) = 9 / 3 = 3$;	$E(6) = 6 / 3 = 2$

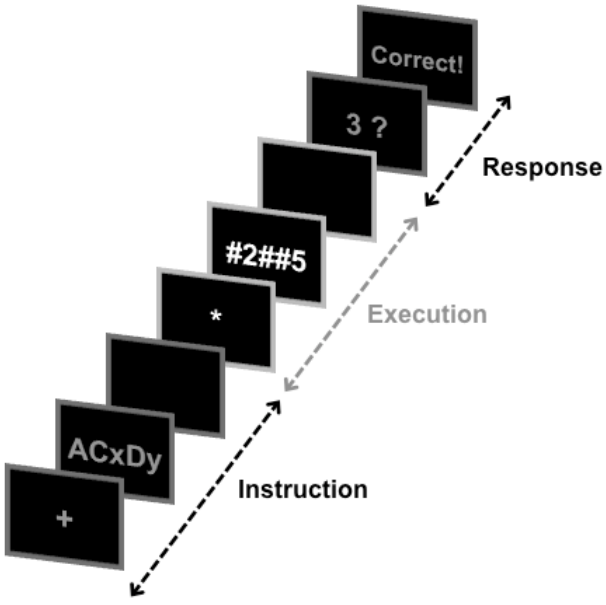

Figure 11: Structure of a sample trial in the experiment

Models for Interpreting Instruction

To explore the nature of the processes involved in interpreting instructions, we modeled the task in both ACT-R and CONDR.

The ACT-R Model

The ACT-R model was designed to execute the entire task, including visually parsing the screen and performing simulated motor responses. During the instruction phase, the model encodes each problem as a series of three consecutive steps. Each step is created by scanning the instruction string right to left, recursively finding the first unattended letter; retrieving the associated operation; and determining whether to apply the operation to either *x*, *y*, or both. During the execution phase, the model simply retrieves the three steps in order, executing the corresponding operations and updating the values of *x* and *y* at the conclusion of each step.

In ACT-R, all the task information must be either available in the buffers or retrieved prior to being used. Thus, some choices had to be made on how to distribute the relevant task information. These choices are usually constrained both by the specific computations available in a module and its established mapping to a brain region (Anderson, 2007). For instance, the intermediate values of x and y, together with the current step's position in the series, were stored in a chunk in the imaginal buffer. This is consistent with the imaginal buffer's association with the parietal cortex, a brain region critically involved in visuo-spatial working memory and mathematical cognition (Anderson, 2005; 2007; Anderson et al., 2008).

The two most critical parts of the model are the chunks representing the problem steps and the production rules that interpret them. Problem steps were maintained in a special module that mimics the computations of the existing goal module. A new module was created because the goal module is associated with internal control states and not with declarative templates for future actions (e.g., Anderson, 2005; Fincham & Anderson, 2006; Anderson et al., 2008). No established association exists between this novel module that processes instructions and a brain region, but some speculations are possible. Its role in holding higher-level representations that tie together lower-level actions suggests an association with the anterior prefrontal cortex (aPFC), which has been often associated with similar functions (Ferrer, O'Hare, & Bunge, 2009).

The model's second key component is the production rules that interpret instructions. These rules differ from standard ACT-R rules in that they use variables to indicate *slot names*, and not only *slot values*. This procedure is needed to properly instantiate operations are referring to either *x* or *y*. The execution of production rules has been associated with the basal ganglia, and basal ganglia activity has been successfully predicted either simply counting the number of production rules fired per time unit (Anderson, 2005; 2007), or by counting the number of variable bindings per time unit (Stocco & Anderson, 2008). Thus, the model predicts that the activity of the basal ganglia should reflect the increased number of variables in the Execution phase. Results of the experimental verification of this production will be presented in Part II.

The CONDR Model

The ACT-R model provides only indirect evidence of the neural basis of interpreting instructions. More compelling evidence can be obtained by modeling the process of following instructions within a framework that directly deals with the underlying biological circuits. The CONDR model was chosen because of its direct connection to ACT-R.

Instructions and the Control of Variable Binding

The very structure of the model suggests one natural way of interpreting instructions. In the routing model, the execution of an operation simply consists in the proper transfer of signals between cortical regions. For example, updating the values of *x* and *y* after an operation consists in copying the representation held in the prefrontal region that retrieves arithmetic facts to the cortical region that temporarily holds either *x* or *y*. This transfer is directed by the proper activation

of cells in the striatum. In fact, any internal operation can be properly represented as a switchboard matrix that shares the same organization of the striatum.

Following this logic, we expanded the routing model by adding a novel cortical area that shares the switchboard organization of **M**, so that variable bindings in the striatal matrix can be properly controlled by the activation of the corresponding cells in the region. In addition to having a switchboard organization, neurons in this region need to have a very low tonic activity; this is required so that their expected activation value $E(x)$ is low, minimizing the effect in calculating the thresholds in Equation (1) and making it easy to bring the activation of projection neurons above the threshold θ. In fact, we ran a number of simulations showing that this mechanism is sufficient to make the model execute arbitrary operations such as the instructed arithmetic operations required by the task.

One can wonder about the biological plausibility of such a hypothetical region. In fact, the anterior part of the prefrontal cortex (aPFC), and in particular the frontal pole, possesses exactly the necessary computational characteristics. Specifically, the aPFC receives massive projections from the frontal lobe, and these projections are topologically organized, thus providing an organization that resembles the frontal projections to the striatum. Also, this region is usually silent during the execution of most tasks, with its most polar part actually deactivates during a task (Gilbert et al., 2006), thus satisfying the condition of a low expected value. Finally, its projections seem to innervate a large part of the head of the caudate nucleus, the most frontal part of the basal ganglia (Di Martino et al., 2008).

Summary

One crucial feature of robust behavior is the capability of interpreting abstract representations that can encode instructions, plans of actions, or intentions. As part of our project, we have developed two different models in two different frameworks that can perform a complex instructed task. The models can also be used to generate neuroimaging predictions, which are described at the end of Part II.

Part II: Experimental Research on Robust Decision-Making

Experimentally, we explored the interactions between the basic functions that underlie robust decision-making. We have run two studies investigating the role of working memory and reward-based learning in sequential decision-making. The crucial problem in sequential decision-making is credit assignment, i.e. the correct attributions of a reward among the different decisions made in a series. While credit assignment for atomic actions is well understood, assignment for sequential actions still constitutes a problem. Our experiments tested two alternative hypothesis: (a) that the connection between current rewards and past actions is mediated by representation in working memory, and (b) that reward is automatically spread, with a temporal discount, among recent actions. We replicated the findings of a previous experiment by Fu and Anderson (2008), while collecting measures of individual differences in working memory capacity and implicit decision-making. Our findings show that experimental manipulations that facilitate working memory improve the quality of decisions. Nonetheless, task performance was not correlated with working memory capacity. The lack of correlation suggests the existence of multiple systems that participants might use to properly credit their decisions. The effect of the manipulation, however, suggests that participants can dynamically exploit features of the task to change the balance between these systems.

Experiment 1

Materials and Methods

Experiment 1 is a follow-up on the experiment by Fu and Anderson (2008). Participants performed a series of 400 trials, each consisting of two decision-making steps. Each decision consisted in choosing between two color names. In each pair, one colors was always associated with high reward probability (80%) and one with low reward probability (20%). Colors occurring in the first choice always had the same reward probability. Instead, reward probabilities associated with colors in the second choice set were dependent on the first choice set. For example, “Blue” could be a high-reward option after choosing between “Yellow” and “Green”, and a low-reward option after “Red” and “Grey”. This made the task non-trivial.

Two factors were manipulated: (a) The Intermediate Credit, i.e., whether the feedback was given only at the end, or intermediate feedback were given after the two choices; and (b) The Intermediate context, i.e., whether the previous choice set was visible or not when making the second decision.

Participants

Twenty-two participants have been run so far. Two participants (#9 and #20) did not complete the experiment, and their partial data were discarded.

Data Analysis

All trials whose latency was < 200ms were discarded from the analysis. Performance was always measured as the proportion of “optimal” choices over a predefined interval. Since these proportions are not distributed normally, they were arcsin-root transformed before being submitted to analysis. This measure will be referred to as P' in the text. For similar reasons, decision latencies were log-transformed before the analysis.

Reading Span Distribution

To make sure that the four groups did not differ on their average Reading Span score, each group’s Reading Span was compared against all the others using a Wilcoxon test. None of the groups differed significantly from the other three [$W(5) > 5$, $p > 0.13$, uncorrected].

Performance by Condition

It is useful to look at the average performance across the entire experiment by condition. The corresponding data were analyzed with a 2x2x2 ANOVA, using Intermediate Credit (True vs False) and Intermediate Context (True vs. False) as between-subjects factors, and and Choice (First vs. Second) as a within-subjects variable. Figure 12 illustrates the effects, and Table 2 reports the results of the analysis.

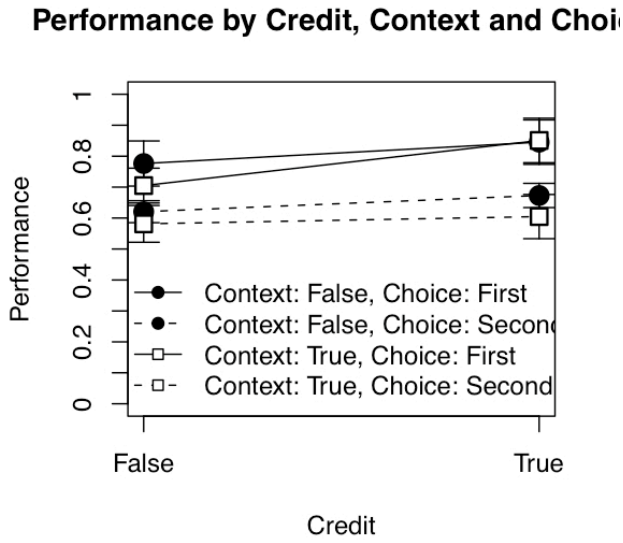

Figure 12: Effects of Credit and Context on performance, grouped by Choice

Table 2: Effects of Credit, Context, and Choice on performance

	DF	*SSE*	*MSE*	*F*	*P*
Context	1, 16	0.02118	0.02118	0.6540	0.4305
Credit	1, 16	0.09477	0.09477	2.9262	0.1065
Context by Credit	1, 16	0.00275	0.00275	0.0850	0.7743
Choice	1, 16	0.45167	0.45167	45.1702	4.913e-06***
Context by Choice	1, 16	0.00040	0.00040	0.0397	0.84460
Credit by Choice	1, 16	0.03128	0.03128	3.1286	0.09599 .
Context by Credit by Choice	1, 16	0.00949	0.00949	0.9487	0.34455

This analysis shows a significant effect of Choice. Performance on the second choice is always *lower* than in the first choice. However, the second choice it is not anymore at chance level. In addition, the analysis shows a marginally significant effect of Credit, with performance being slightly higher when the Intermediate Credit was given (P' = 1.06) than when it was not (P' = 0.97) and a marginally significant interaction between Credit and Choice.

First and Second Choice by Participant

It is interesting to look at the individual data. Figure 13 shows the difference between First and Second choice in all participants. You can refer to Table 1 to check out each participant's condition.

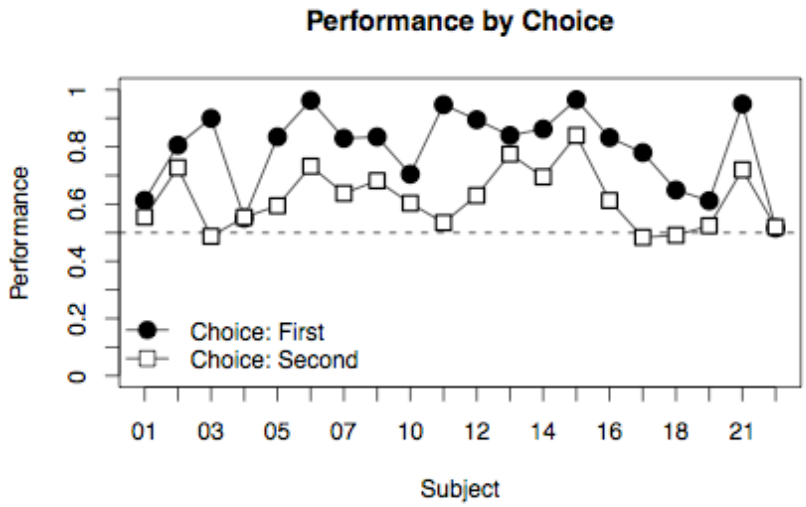

Figure 13: First and second choice performance by participant

Besides three of them (#1, #4 and #22, who show almost random performance), participants do show some effects of learning in the second choice. Second-choice performance is always inferior to the First choice. Nonetheless, performances in the two choices are significantly correlated [$r = 0.58$, $t(18) = 0.008$].

Analysis by Block

To have a better idea of learning, we can look as performance changes through the experimental blocks. The experimental trials were divided into 8 blocks of 50 trials each.

It is difficult to visualize the interactions between Block, Choice, Context, and Credit. Therefore, I broke it down into four different figures. Figure 14 shows the effects of Block and Choice on Performance, divided by Context. Figure 15 shows the effect of Block and Choice on Performance, divided by Credit. Tables 3 and 4 report the results of the corresponding ANOVAs.

Performance by Block, Choice, and Context

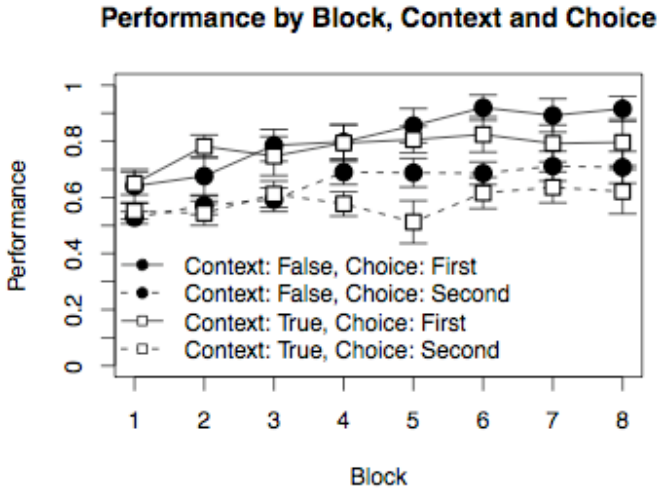

Figure 14: Interaction of Block and Choice on performance, grouped by Context type

Table 3: Effects of Context, Choice, and Block on P'

	DF	*SSE*	*MSE*	*F*	*P*
Context	1, 18	0.3638	0.3638	1.013	0.3275
Choice	1, 18	4.8944	4.8944	50.5353	1.262e-06 ***
Context by Choice	1, 18	6.589e-06	6.589e-06	0.0001	0.9935
Block	7, 126	2.1271	0.3039	9.7895	1.054e-09 ***
Block by Context	7, 126	0.4127	0.0590	1.8996	0.07487 .
Block by Choice	7, 126	0.35044	0.05006	3.7488	0.00101 **
Block by Context by Choice	7, 126	0.21581	0.03083	2.3086	0.03008 *

Not surprisingly, the Choice and Block are both significant. The main effect of Context is not, but the Context interacts significantly with Choice, Block, and with their interaction. This means that the context affects the learning rate (its effect on the Block factor), and the difference between the two choices (its effect on the Choice factor). Interesting, the presence of a Context seems to hinder, instead of helping, performance.

Performance by Block, Choice and Credit

As we know from the previous analysis, the main effect of Credit is marginally significant. It is interesting to notice that the Credit interacts with Block (i.e., facilitates learning for the first choice) but does not seem to interact with the effect of Choice, or with the Choice by Block interaction (although the interaction with Choice has a low *p*-level).

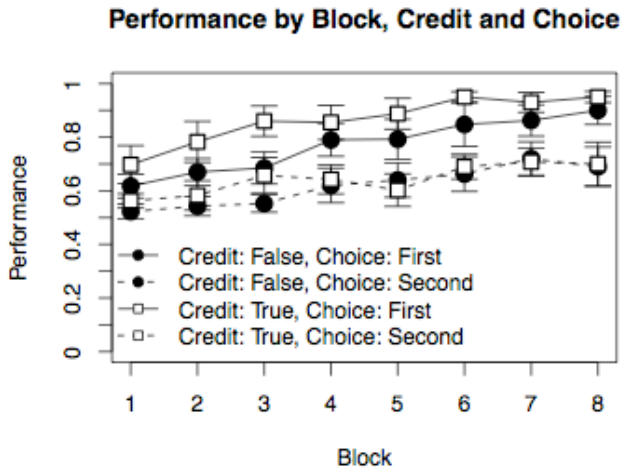

Figure 15: Interaction of Block and Choice on Performance, grouped by Credit

Table 4: Effects of Credit, Choice, and Block on P'

	DF	*SSE*	*MSE*	*F*	*P*
Credit	1, 18	0.9628	0.9628	2.9551	0.1028
Choice	1, 18	4.8944	4.8944	58.1077	4.841e-07***
Credit by Choice	1, 18	0.2272	0.2272	2.6973	0.1179
Block	7, 126	2.1271	0.3039	9.0223	5.41e-09***
Block by Credit	7, 126	0.0802	0.0115	0.3400	0.9341

Block by Choice	7, 126	0.35044	0.05006	3.4083	0.002278**
Block by Credit by Choice	7, 126	0.04775	0.00682	0.4644	0.858702

Correlations with Reading Span

We did perform a simple individual difference analysis by correlating participants' performance with their Reading Span score. Because performance on the first and on the second choice are so different, two separate correlation analysis were performed on these two measures. Table 5 reports the results

Table 5: Correlations between Reading Span and Performance

Analysis	*R*	*DF*	*T*	*P*
Reading span with *P'*	0.07978134	18	0.3395660	0.3690
Reading span with *P'* in the First Choice	0.1175037	18	0.5128226	0.3071
Reading span with *P'* in the Second Choice	0.0301295	18	0.1273365	0.4500

There does not seem to be any significant correlation between Reading Span and Performance (as measure by *P'*) for either choice.

Summary

Experiment 1 provided evidence for the importance of intermediate credit assignment in decision-making. The existence of intermediate assignment, however, was not sufficient to level the difference between the two choices, hinting at the existence of powerful order effects and at the difficulty of decision-making when complex dependencies exists between choices. Furthermore, making the choice context explicit, which was supposed to highlight the dependency, had the counter-intuitive effect of harming performance.

Experiment 2: Encoding of Action vs. Context

Rationale

In the first experiment, the presence of an Intermediate Credit was found to improve participant's performance. Contrary to our expectations, the presence of an Intermediate Context was found to have a slightly detrimental effect. That is, when participants were reminded of the options available during the first choice, they tend to perform worse.

One potential explanation for this inconsistency was that the type of Context provided was somewhat misleading. In Experiment 1, Context was provided by reminding participants of the options available in the first choice (i.e., "Blue – Yellow"). Some neuroimaging evidences (e.g., Tricomi, Delgado, & Fiez, 2004) suggest that reward-related learning is mediated by direct action contingency. That is, rewards are tied to representations of *actions*, and not simply to their *context*. To test this hypothesis, we decided to modify the context type so that it contains only the color participants had actually selected in the first choice.

Materials and Methods

The experiment used exactly the same materials as Experiment 1. Only, this time the *context* was changed to the color actually picked by the participants in the first choice (e.g., "Yellow", instead of "Blue – Yellow").

Participants

Thirteen participants were recruited (they are going to be fourteen soon). One participant (#25)

was run in the wrong condition, and her data have been discarded from the analysis. Data from one other participant (#24) were corrupted while being transferred from the local computer to the data server.

Instead of running a replica of the entire experiment, only the two conditions that included a Context were run. In both conditions, the Context was always present. Additionally, the Context always consisted of the choice the participants had selected. The Intermediate Credit was present in one condition ($N = 5$ participants) and absent in the other ($N = 6$ participants).

As in Experiment 1, a performance index P was calculated as the arcosine of the square root of the proportion of correct choices. As in Experiment 1, a choice was considered correct when they it corresponded to the color with the highest reward probability, independent of the reward actually being delivered or not. Finally, trials where the response latency was < 200ms were excluded from the analysis.

Effect of Credit

As a first check of our data, we assessed whether the Credit factor had any effect on performance. Although the effects were not significant yet [$t(6) = 1.24$, $p = 0.25$], the manipulation was going in the right direction. Participants who did not see the Intermediate Credit performed worse ($P = 1.02$, $SD = 014$) than those who did ($P = 1.15$, $SD = 0.18$)

Effect of Context Type

We compared data from this experiment against data from Experiment 1. A first interesting analysis is to compare the effects of two types of Intermediate Context, i.e., "Choices" (as in Experiment 2) or "Options" (as in Experiment 1). To perform this analysis, data from Experiment 2 were pooled together with the two conditions in Experiment 1 where participants were provided with an Intermediate Context. This way, the Context was always present in this pool of participants, and the only factors manipulated were the type of Context and the presence of Intermediate Credit. Figure 16 illustrates the Effect of Credit and Context Type on Performance for the two choices. Table 6 reports the results of the corresponding ANOVA.

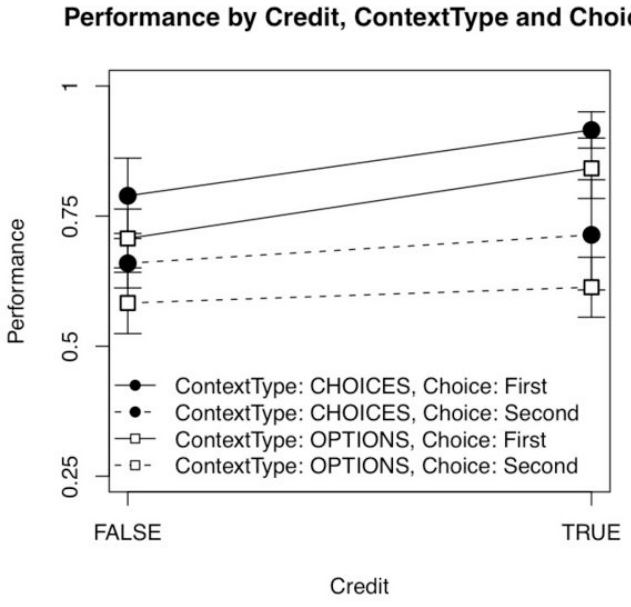

Figure 16: Effects of Credit, Choice, and intermediate Context type on performance

Table 6: Effects of Credit, Choice, and Context type on performance

Factor	*DF*	*SSE*	*MSE*	*F*	*p*
Context Type	1, 18	0.09682	0.09682	2.3584	0.14200
Credit	1, 18	0.15400	0.15400	3.7513	0.06862
Context Type by Credit	1, 18	0.00142	0.00142	0.0345	0.85466
Choice	1, 18	0.49161	0.49161	50.2201	1.317e-06***
Context Type by Choice	1, 18	0.00006	0.00006	0.0057	0.94068
Credit by Choice	1, 18	0.04473	0.04473	4.5698	0.04651*
Context Type by Credit by Choice	1, 18	0.00112	0.00112	0.1147	0.73877

Discussion

In summary, it seems clear that the Context type *does* affect performance. In particular, and as expected, when the context is represented as the previous choice (i.e., "Blue") instead of the previous options (i.e., "Blue - Yellow" or "Yellow - Blue"), participants tend to do better, in both the first and the second choice.

Re-evaluation of Experiment 1

We can now go back to Experiment 1 and simply substitute the data from the old participants in the two Context condition with the new data from Experiment 2. This will permit an analysis of both Credit and Context factors, but with Context now being the presence of the chosen color (instead of the previous options).

Main Effects of Context and Credit on Performance

As in the Experiment 1, we started by looking at the effects of Credit and Context on the raw performance for First and Second choice. The results are very similar to the homologous results in Experiment 1; however, this time the presence of an Intermediate Context is improving, and not harming, participants' performance. The results are reported in Figure 17 and Table 7.

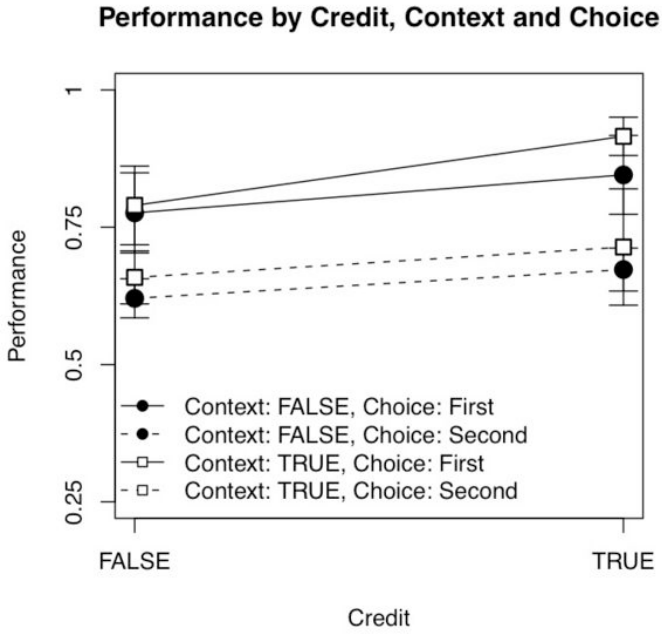

Figure 17: Effects of Context and Credit on performance, grouped by choice

Table 7: Effects of Context, Credit and Choice on performance

Factor	*DF*	*SSE*	*MSE*	*F*	*p*
Context	1, 17	0.02773	0.02773	0.6723	0.4236
Credit	1, 17	0.11871	0.11871	2.8784	0.1080
Context by Credit	1, 17	0.00638	0.00638	0.1547	0.6989
Choice	1, 17	0.45497	0.45497	56.9820	7.974e-07***
Context by Choice	1, 17	0.00004	0.00004	0.0045	0.9474
Credit by Choice	1, 17	0.01653	0.01653	2.0702	0.1684
Context by Credit by Choice	1, 17	0.00205	0.00205	0.2569	0.6188

Experiment 3: Testing Models for Interpreting Instructions

As part of our experimental investigations, we obtained an award that allowed us to run a pilot neuroimaging study. This study was aimed at identifying the neural correlates of adaptive behavior. It adopts a novel paradigm where planned changes in behavior can be separated from their subsequent execution, thus permitting for the first time to isolate the two corresponding networks or regions and their connections.

Ten participants were recruited to perform the task previously described while lying in a 3T fMRI scanner. Their brain activity was recorded at a rate of a full volume acquisition every 2 seconds, with 34 oblique slices acquired for each volume. Each participant solved 80 problems, divided into four blocks of 20 trials each. Unlike most fMRI experiments, each problem was self-paced.

In addition to the distinction between encoding and executing a set of instructions, the experiment manipulated the amount of practice as a second factor. This manipulation provides an additional means to isolate the specific act of interpreting instructions, which is important when analyzing data with a limited number of participants (see below). Practice was manipulated by having participants perform a subset of the problems before the experiment. During the experiment, half of trials were novel and half came from the subset of practiced trials.

Results

Data for the experiment were used to test the predictions of two models we developed that could perform the task (see above). Because the low number of participants limited the statistical power of traditional analysis, we performed a conjunction analysis, using statistical parameter maps thresholded at a liberal voxel-level value ($p < 0.01$, uncorrected) to isolate regions that are activated in two or more target contrasts.

The ACT-R model predicts that the module corresponding to the aPFC region should be more active in Novel than Practiced trials, in both the Instruction and Execution phases. Thus, we created to statistical parameter maps (one for the Instruction phase, one for the Execution phase) that identified those voxels that were statistically more active during the Novel than during the Practiced trials (i.e., Novel > Practiced). As predicted, the analysis identified a cluster of voxels located in the aPFC region, with smaller cluster located in even anterior position in the frontal lobe. The results of this analysis are illustrated in the top part of Figure 18; the crosshairs highlight the aPFC regions.

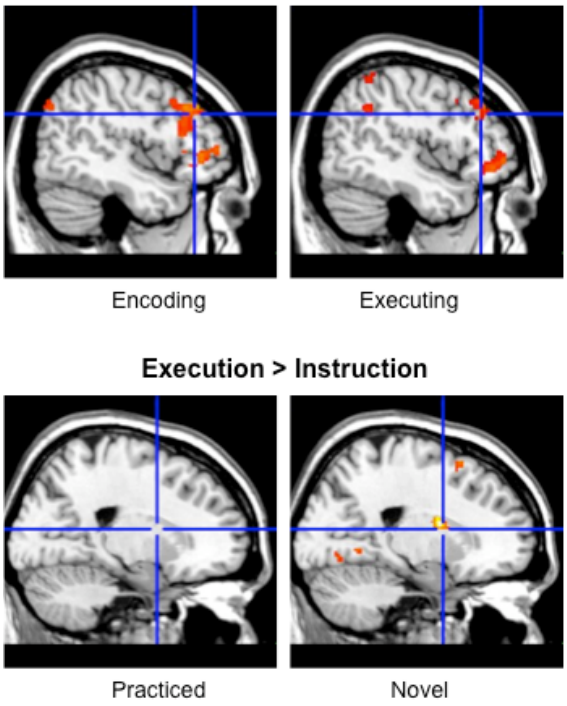

Figure 18: Results of Experiment 3

Both the ACT-R and the conditional routing model predict that the basal ganglia should be more active during the Execution phase than during the Instruction phase. Additionally, both models predict that this asymmetry should hold for Novel problems only; Practiced problems can be executed as a routine, without referring to the original instructions, and there is no reason to expect any additional basal ganglia involvement during their execution. To verify this hypothesis, we created two new contrast maps that identify those voxels more active in the Execution than the Instruction phase (i.e., Execution > Instruction) in the Novel and in the Practiced problems, respectively. As predicted, we found one cluster of voxels that was more active during the Execution phase and corresponded to the right striatum; it is indicated by the crosshairs in the bottom part of Figure 18. As predicted this cluster showed up only in the contrast map obtained from Novel trials; Practiced problems did not show, in fact, any voxel that was more active during the Execution phase. In summary, our preliminary results support our models' predictions and permit to identify two regions crucially involved in interpreting instructions: the aPFC, probably responsible for encoding and accessing abstract representations of cognitive actions, and the basal ganglia, probably responsible for performing the necessary variable bindings while interpreting instructions.

References

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences, 12(10)*, 366-375.

Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science, 29(3)*, 313-341.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.

Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing problem solving in real time: fMRI analysis of the subject-paced Tower of Hanoi. *Journal of Cognitive Neuroscience, 17(8)*, 1261-1274.

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, 12, 136-143.

Anderson, J. R. & Gluck, K. (2001). What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr & S. M. Carver (Eds.) *Cognition & Instruction: Twenty-five years of progress*, 227-262. Erlbaum.

Anderson, J. R., Taatgen, N. A., & Byrne, M. D. (2005). Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance, 31(4)*, 749-761.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review, 114(3)*, 632-656.

Barto, A. G. (1995). Adaptive critics and the basal ganglia. In J. C. Houk, J. L. Davis & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 215-232). Cambridge, MA: MIT Press.

Best, B. J. & Lebiere, C. (2006). Cognitive agents interacting in real and virtual worlds. In R. Sun (ed.), *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*. Cambridge University Press; New York, NY, 186-218.

Bolam, J. P., Hanley, J. J., Booth, P. A., & Bevan, M. D. (2000). Synaptic organisation of the basal ganglia. *Journal of Anatomy, 196(4)*, 527-542.

Braver, T.S., and Cohen, J.D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In Monsell, S. and Driver, J. (Eds.) *Attention and Performance XVIII*. (pp. 713-737). Cambridge, MA: MIT Press.

Brown, T. H., Kairiss, E. W., & Keenan, C. L. (1990). Hebbian synapses: Biophysical mechanisms and algorithms. *Annual Review of Neuroscience, 13*, 475-511.

Byrne, M. D., & Kirlik, A. (2005). Using computational cognitive modeling to diagnose possible sources of aviation error. *International Journal of Aviation Psychology, 15*, 135-155.

Calabresi, P., Gubellini, P., Centonze, D., Picconi, B., Bernardi, G., Chergui, K., et al. (2000). Dopamine and cAMP-regulated phosphoprotein 32 kDa controls both striatal long-term depression and long-term potentiation, opposing forms of synaptic plasticity. *Journal of Neuroscience, 20(22)*, 8443-8451.

Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Brain Research—Cognitive Brain Research, 25(3)*, 607-623.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science, 210(4466)*, 207-210.

Danker, J. F., Gunn, P. & Anderson, J. R. (2008). A rational account of memory predicts left prefrontal activation during controlled retrieval. *Cerebral Cortex*, *18*, 2674-2685.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., Biswal, B., Walters, J. R., Castellanos, F. X., & Milham, M. P. (2008). Functional connectivity of human striatum: A resting state fMRI Study. *Cerebral Cortex, 18*, 2735-2747.

Ferrer, E., O'Hare, E. O., & Bunge, S. A. (2009). Fluid reasoning and the developing brain. *Frontiers in Neuroscience*, *3(1)*, 46-51.

Fincham, J. M. & Anderson, J. R. (2006). Distinct roles of the anterior cingulate and prefrontal cortex in acquisition and performance of a cognitive skill. *Proceedings of the National Acedemy of Sciences*, *103*, 12941-12946.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective & Behavioral Neuroscience*, *1(2)*, 137-160.

Fu, W-T. & Anderson, J. R. (2004). Extending the Computational Abilities of the Procedural Learning Mechanism in ACT-R. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 416-421). August 4-7, Chicago, USA.

Fu, W.-T., Anderson, J. (2008), Solving the Credit Assignment Problem: Explicit and Implicit Learning of Action Sequences with Probabilistic Outcomes. *Psychological Research, 72 (3)*, 321-330.

Gerstner, W., & Kistler, W. M. (2002). Mathematical formulations of Hebbian learning. *Biological Cybernetics, 87(5-6)*, 404-415.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (Area 10): A meta-analysis. *Journal of Cognitive Neuroscience*, *18*, 932–948.

Graveland, G. A., Williams, R. S., & DiFiglia, M. (1985). A Golgi study of the human neostriatum: neurons and afferent fibers. *Journal of Comparative Neurology, 234(3)*, 317-333.

Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy, 26(4)*, 317-330.

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *Journal of Neuroscience, 20(6)*, 2369-2382.

Hajos, M., & Greenfield, S. A. (1994). Synaptic connections between pars compacta and pars reticulata neurones: electrophysiological evidence for functional modules within the substantia nigra. *Brain Research, 660(2)*, 216-224.

Hattori, T., Fibiger, H. C., McGeer, P. L. (1975). Demonstration of a pallido-nigral projection innervating dopaminergic neurons. *The Journal of Comparative Neurology, 162*, 487–504.

Hawkins, J., & Blakeslee, S. (2004). *On intelligence*. New York, NY: Times Books.

Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance, 28(3)*, 527-545.

Hill, N. M., & Schneider, W. (2006) Brain changes in the development of expertise: Neurological evidence on skill-based adaptations. In K. A. Ericsson, N. Charness, P. Feltovich, and R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance*. New York:

Cambridge University, 653-682.

Houk, J. C. (2005). Agents of the mind. *Biological Cybernetics, 92(6)*, 427-437.

Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks, 15(4-6)*, 535-547.

Just, M. A., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Cognitive, Affective and Behavioral Neuroscience, 7(3)*, 153-191.

Laird, J. E. (2008). Extending the Soar Cognitive Architecture. *Proceedings of the First AGI Conference*, 224-235

Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986). Chunking in soar: The anatomy of a general learning mechanism. *Machine Learning, 1(1)*, 11-46.

McClelland, J.L., McNaughton, B.L. & O'Reilly, R.C. (1995). Why There are Complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connectionist Models of Learning and Memory. *Psychological Review, 102*, 419-457.

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience, 11(1)*, 103-107.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance. 1. Basic mechanisms. *Psychological Review, 104(1)*, 3-65.

Miller, J. D., Sanghera, M. K., & German, D. C. (1981). Mesencephalic dopaminergic unit activity in the behaviorally conditioned rat. *Life Sciences, 29*, 1255-1263.

Newell, A. N. (1973). Production systems: Models of control structure In W. G. Chase (Ed.), *Visual information processing* (pp. 526-547). New York, NY: Academic Press.

Nicola, S. M., Surmeier, J., & Malenka, R. C. (2000). Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annual Review of Neuroscience, 23*, 185-215.

O'Reilly, R.C. (1996). Biologically Plausible Error-driven Learning using Local Activation Differences: The Generalized Recirculation Algorithm. *Neural Computation, 8*, 895-938.

O'Reilly, R.C. (1998). Six Principles for Biologically-Based Computational Models of Cortical Cognition. *Trends in Cognitive Sciences*, 2, 455-462.

O'Reilly, R.C. (2001). Generalization in Interactive Networks: The Benefits of Inhibitory Competition and Hebbian Learning. *Neural Computation*, *13*, 1199-1242.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283-328.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience*. Cambridge, MA: MIT Press.

O'Reilly, R.C. & Norman, K.A. (2002). Hippocampal and Neocortical Contributions to Memory: Advances in the Complementary Learning Systems Framework. *Trends in Cognitive Sciences, 6*, 505-510.

O'Reilly, R.C. & Rudy, J.W. (2001). Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review*, *108*, 311-345.

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.

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience, 25*, 563-593.

Parthasarathy, H. B., Schall, J. D., & Graybiel, A. M. (1992). Distributed but convergent ordering of corticostriatal projections: analysis of the frontal eye field and the supplementary eye field in the macaque monkey. *Journal of Neuroscience, 12(11)*, 4468-4488.

Qin, Y., Sohn, M-H, Anderson, J. R., Stenger, V. A., Fissell, K., Goode, A., & Carter, C. S. (2003). Predicting the practice effects on the blood oxygenation level-dependent (BOLD) function of fMRI in a symbolic manipulation task. *Proceedings of the National Academy of Sciences, 100*, 4951-4956.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M., Pardo, J. V., Fox, P. T., et al. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex, 4(1)*, 8-26.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors, 48*, 362-380.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80(1)*, 1-27.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241-263.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., et al. (2001). Virtually perfect time sharing in dual-task performance: uncorking the central cognitive bottleneck. *Psychological Science, 12(2)*, 101-108.

Smith, A. D., Bolam, J. P. (1990). The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurones. *Trends in Neurosciences. 13(7)*, 259-65.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence, 46(1-2)*, 159-216.

Sohn, M.-H., Albert, M. V.., Stenger, V. A, Jung, K.-J., Carter, C. S., & Anderson, J. R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proceedings of National Academy of Science, 104*, 10330-10334.

Sohn, M.-H., Goode, A., Stenger, V. A, Carter, C. S., & Anderson, J. R. (2003). Competition and representation during memory retrieval: Roles of the prefrontal cortex and the posterior parietal cortex, *Proceedings of National Academy of Sciences, 100*, 7412-7417.

Stewart, T. C., & Eliasmith, C. (2008). Building Production Systems with Realistic Spiking Neurons. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*, Washington, D.C.

Stocco, A., & Anderson, J. R. (2008). Endogenous control and task representation: An fMRI study in algebraic problem-solving. *Journal of Cognitive Neuroscience*, *20(7)*, 1300-1314.

Stoffers, D., Bosboom, J. L., Deijen, J. B., Wolters, E., Stam, C. J., & Berendse, H. W. (2008). Increased cortico-cortical functional connectivity in early-stage Parkinson's disease: An MEG study. *Neuroimage, 41(2)*, 212-222.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning, 3(1)*, 9-44.

Taatgen, N. A., Huss, D., Dickison, D. & Anderson, J. R. (2008). The acquisition of robust and flexible cognitive skills. *Journal of Experimental Psychology: General, 137(3)*, 548-565.

Taatgen, N. A., & Lee, F. J. (2003). Production compilation: A simple mechanism to model complex skill acquisition. *Human Factors, 45(1)*, 61-76.

Tepper, J. M., & Bolam, J. P. (2004). Functional diversity and specificity of neostriatal interneurons. *Current Opinion in Neurobiology, 14(6)*, 685-692.

Tepper, J. M., Martin, L. P., & Anderson, D. R. (1995). $GABA_A$ receptor-mediated inhibition of rat substantia nigra dopaminergic neurons by pars reticulata projection neurons. *Journal of Neuroscience, 15(4)*, 3092-3103.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K. & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences USA*, 94(26), 14792–14797.

Touretzky, D., & Hinton, G. H. (1988). A distributed connectionist production system. *Cognitive Science, 12*, 423-466.

van der Velde, F., & de Kamps, M. 2006. Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences, 29(1)*, 37-70.

Wickens, J. (1997). Basal ganglia: structure and computations. *Network: Computation in Neural Systems*, 8, R77-R109.

Wickens, J. R., Begg, A. J., & Arbuthnott, G. W. (1996). Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience, 70(1)*, 1-5.

Appendix A: Implementation of the CONDR Model

This appendix will give an overview of the model's different types of units, their connections, and functions.

Model Neurons

In the model, neurons were implemented as simple computational units that apply an activation function f over an input value η to yield an activation value, denoted by x. The input value η is simply the sum of all the activations coming from the projecting neurons, weighted by the corresponding synaptic strengths:

$$\eta = \sum_j w_j x_j$$

where w_j is the value (or "synaptic weight") of the synapse from neuron j, x_j is the activation of neuron j. This is perhaps the simplest and most common representation for artificial neurons, and is widely adopted in many biological models (see Rolls &Treves, 1998; O'Reilly & Munakata, 2000). The activation value x is obtained from the net input η by applying the activation function f:

$$x = f(\eta - \theta)$$

where θ is the neuron's threshold, which can be thought of an initial resistance of every neuron to be excited. A negative threshold (so that the quantity $\eta - \theta$ is positive in absence of direct stimulation) can be used to model neurons with high baseline activities, or to compensate the effects of convergent inhibitory projections. The activation value x is supposed to be the computational counterpart of a neuron's firing rate. Note that a neuron's dynamic is completely characterized by its activation function and threshold.

With the exception of striatal interneurons (discussed below), all the neurons in the model use the hyperbolic tangent as their activation function:

$$x = \tanh(\gamma [\eta - \theta]_+)$$

where γ is the gain parameter that determines the curves' steepness, and the $[x]_+$ notation indicates that negative values of x are treated as zeroes. This ensures that the output of the function is in the range [0, 1]. Together with the sigmoid function, the hyperbolic tangent is among the simplest formulae that fit the change of spiking rates following changes in membrane potential in biological neurons; the curve also closely mimics the variation of spike rates to a change in the membrane potentials in biological neurons (see O'Reilly & Munakata, 2000). Table A1 details the values of γ and θ for each type of neuron in the model. Figure A1 gives a visual rendition of the corresponding activation curves.

Special Activation Function for Striatal Interneurons

Striatal interneurons exhibit special dynamics. They are tonically active and exert inhibitory pressure on striatal projection neurons, unless cortical activation reduces their firing rates. This behavior is likely produced by the interaction between cholinergic and GABA-ergic interneurons (e.g., Tepper & Bolam, 2004; see Figure 2). To account for this behavior, the only type of striatal interneurons in our model were provided with a special activation function, consisting of a sigmoid

function with positive exponent. This function is monotonically decreasing, so that increased cortical inputs decrease the activity of interneurons. The function and its parameters are reported in Table A1, and visually depicted in Figure A1.

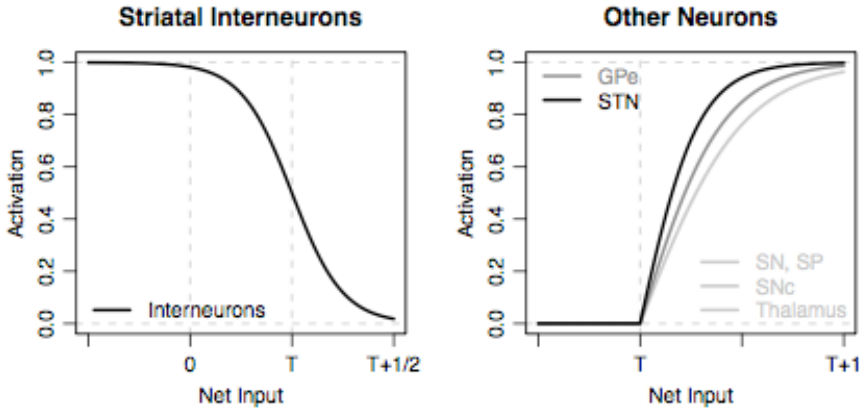

Figure A1: A visual rendition of the activation functions used in the model. Left: the monotonically decreasing function used to simulate the firing rate patterns of striatal interneurons. Right: The monotonically increasing functions used to simulate the firing patterns of all the other neurons.

Baseline Activation Values

Each neuron also has an associated quantity called baseline activation value, which is indicated as $\langle x \rangle$. The baseline can be interpreted as the neuron's tonic activity. The baseline value provides a simple means to measure how much of a certain neuron's activation is due to its current inputs. It plays an important role in the Hebbian learning rule that is used in the model:

$$\Delta w_{i,j} \approx r\,(x_i - \langle x_i \rangle)\,(x_j - \langle x_j \rangle)$$

In this rule, subtracting the baseline from the activation prevents two neurons to become strongly associated when their activity is due to their "usual", tonic condition. It also makes it possible for a synapse to lose strength, whenever activation in one neuron is coupled with a decrease of activation in the other. The baselines for each different neuron type were calculated as their activation values in absence of stimulation, i.e., when $\eta = 0$.

"Up" and "Down" States in Striatal Projection Neurons

Projection neurons in the striatum have special dynamics. They cannot be excited while they are in the "down" state. Cortical activity puts them in an "up" state; when in "up" state, an increase in excitatory input or a decrease of inhibition triggers a response (Wilson, 1993; 1995; Bolam, et al. 2000). A realistically complex model of this behavior was beyond the scope of our research. A simple approximation, however, consists in using neurons with a dynamic threshold. A threshold value θ is said to be dynamic when it is allowed to change over time. Some learning rules that have been found biological support, such as the BCM rule (Bienenstock, Cooper, & Munro, 1982) make use of dynamic thresholds. In the model, the dynamic threshold θp for a projection neuron p approximates the expected input from striatal interneurons when cortical patterns are being gated.

Table A1: Summary of the different types of neurons

Neuron Type	*Activation Function*	*Gain Parameter*	*Threshold θ*
Striatal interneuron	$x = 1 / (1 + e^{\gamma [\eta - \theta]})$	$\gamma = 8.0$	$\theta = 1/2$
Striatal projection neuron (SN/SP)	$x = \tanh(\gamma [\eta - \theta]_+)$	$\gamma = 2.0$	$\theta_p \approx \sum_i w_{p,i} x_i$
Thalamic neuron		$\gamma = 2.0$	$\theta = -1$
SNr/GPi neurons		$\gamma = 2.0$	$\theta = -2$
STN neurons		$\gamma = 3.0$	$\theta = -1$
GPe neuron		$\gamma = 3.5$	$\theta = -1$
SNc interneuron		$\gamma = 2.0$	$\theta = -1$
SNc dopamine neuron		$\gamma = 2.0$	$\theta = -1$

That is, the threshold is adapted to match the amount of inhibition that a projection neuron receives from interneurons when the projection neuron is nonetheless firing:

$$\theta_p \approx \sum_i w_{p,i} x_i^*$$

where x_i^* indicates the average activation of interneuron *i* when cortical signals are allowed to pass. This value depends on the routing patterns encoded in the model, and was therefore calculated separately for each set of simulations. Note that the values of θ_p are dynamic because they depend on the strength of synapses wp,i. Therefore, they are recalculated every time the synaptic weights are changed by Hebbian learning. A model projection neuron's activation remains constant and equal to 0 until the sum of all its inputs stays below θ_p. This corresponds to the "down" state. When the interneuron inhibition matches the threshold, the neuron reaches the "up" state, and any additional input, from either the cortex or interneurons, increases its activation.

More on Synapses

Inhibitory synapses were encoded as negative weights, and excitatory synapses were encoded as positive weights. While the value of the synaptic change was left free to change according to Hebbian learning, no synapse could ever change sign. That is, negative synapses could not rise above zero, and positive synapses could not decrease below zero. This reflects the biological fact that inhibitory synapses cannot turn excitatory, and vice-versa. The only exception to this rule consists in the synapses between cortical neurons and striatal interneurons. The reason for this exception is that striatal interneurons represent the net contribution of GABA-ergic and cholinergic interneurons.

Receptive Fields and Representation Compression

Nuclei in the basal ganglia have increasingly smaller size, which suggests a progressive "funneling" of information (Alexander, DeLong, & Strick, 1986). This is an important characteristic of the basal ganglia physiology, and needs to be addressed in a realistic model of the circuit. The easiest way to model this compression of information is to arrange the synaptic inputs so that a neuron from a smaller region receives inputs from many neurons that occupy the same position in a larger, input nucleus. Let us suppose that the projecting region has m neuron, and its target region contains n neurons (with $n < m$). If we indicate with j a neuron in the projecting region, and with i a neuron in the target region, then the synaptic weight wi,j is given by:

$$w_{i,j} = G(i - j \times (n / m), \sigma)$$

where $G(x, \sigma)$ is a Gaussian (normal) function with mean 0 and standard deviation σ. In the expression, the term n/m is used to the express the position of the neuron j within a range between 0 and n. This way, the relative positions of neurons i and j in the two regions can be compared. The term $i - j \times (n / m)$ can be read as the difference between the two relative positions. When this difference is zero, j is at the center of i's receptive field. Figure A2 illustrates the shape of such receptive field in the case of projections from the cortex (m = 100) to the striatum (n = 10. This m/n ratio is actually close to the ratio of cortical projection neurons to striatal projection neurons, as estimated by Zheng and Wilson, 2002). In the model, similar functions are used to model connections between all nuclei, which usually differ in size. Notice that synaptic weights depend only on m and n and the free parameter σ. In the model, σ = ½ across all projections. The only exception was the striatal projections from interneurons to output neurons, where $\sigma = n/2$. This created an almost uniform inhibitory pressure.

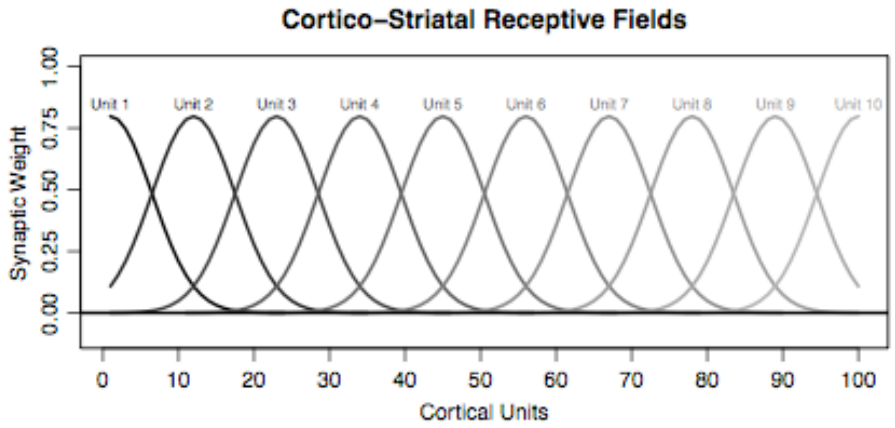

Figure A2: A visual rendition of the Gaussian striatal receptive fields used in the model. In this Figure, ten striatal units receive inputs from 100 cortical units. Their receptive fields are shown as bell curves of different colors in the figure. They are shaped in such a way that each striatal unit is maximally sensitive to those cortical neurons that occupy a similar position in the cortical regions. This way, cortical topology is maintained within striatal subdivision.