J. Intell. Syst. **20** (2011), 373–393
DOI 10.1515/JISYS.2011.019

A Neurocomputational Model for the Relation Between Hunger, Dopamine and Action Rate

Lovekesh Vig, Ashish Gupta and Abhinandan Basu

Abstract. A number of conditioning experiments utilize food as a reward. Hunger is considered to be a critical factor governing the animal's behavior in these experiments. Despite its significance, most theories of animal conditioning fail to take hunger into consideration while analyzing the behavioral data. In this paper, we analyze the neuroscientific data supporting the hypothesis that hunger and food consumption affect the brain's dopamine system, which in turn governs the animal's behavior. According to this hypothesis, chronic hunger results in a decrease in the extra-cellular dopamine levels in the animal's brain. This decrease is believed to trigger a series of reactions that increase the responsivity of the phasic dopamine system. A direct consequence of this is an increased vigor of all dopamine-dependent behaviors. The level of extra-cellular dopamine is also modulated by the process of food consumption via the neurotransmitter Acetylcholine. Food consumption raises the dopamine above the baseline level. This rise depends on the animals' hunger, which when satisfied increases the level of Acetylcholine, which causes the dopamine level to fall back to the baseline. Thus, extra-cellular dopamine governs the response vigor, with an increase in dopamine resulting in a more vigorous response. This paper makes two primary contributions. Firstly, we present an abstract mathematical model based on the above hypothesis. Our mathematical model is able to provide a simple explanation for a number of behavioral findings. Another contribution of this paper is the development of a neurocomputational Leabra model of dopamine and acetylcholine activity in the basal ganglia to incorporate hunger and satiation. The experimental results obtained from this neural model are also largely in agreement with behavioral findings.

Keywords. Hunger, Dopamine, Acetylcholine, Satiation, Leabra.

1 Introduction

Operant conditioning involves the modification of an animal's behavior through the association of some reward. As the association becomes stronger, the frequency with which the reward-related action is executed also increases. Other possible factors governing the action frequency include the need for the reward, fatigue, choice of other possible rewarding actions, availability of other possible rewards, the amount of effort required for executing the action, the magnitude of the reward and the time gap between action execution and reward delivery.

A number of conditioning studies use food as reward. In such studies, hunger is considered to be one of the most critical factors governing behavior. Studies suggest that chronic hunger makes the animal's behavior more responsive to the phasic dopamine activity. This heightened responsivity leads to an increased vigor of all dopamine-dependent behaviors. Further, there is evidence that food consumption results in an increase in the extra-cellular dopamine level in the brain. This increase occurs only if the animal is hungry. Once the animal becomes sated due to continued food consumption, the dopamine level falls back to the baseline level. An interesting finding is that this modulation in the dopamine level seems to govern the food-related action rate.

In this paper, we have explored the behavioral and neuroscientific evidence supporting the above hypothesis. We present a mathematical model derived from our previous work [30] that is based on the above hypothesis. This model provides a simple explanation for a number of puzzling behavioral findings related to animal conditioning. Subsequently a more biologically plausible neurocomputational model is provided that sheds light on the role of actylcholine in signalling satiation. Experimental results are shown to be in qualitative agreement with behavioral findings.

The paper is organized as follows: The next section reviews the behavioral and neuroscientific data describing the relation between hunger, food consumption, dopamine level and action rate. Subsequent sections propose our mathematical and neurocomputational models followed by simulation results. The final section offers description of the future challenges.

2 Background

Hunger is seen to affect the action rate in two different ways. The first deals with the effects of chronic hunger. In this case animals are typically maintained below the baseline body weight for the duration of the conditioning studies. These studies can last from a few weeks to a few months period. It is observed that food deprived animals show lower levels of extra-cellular dopamine as compared to the control animals [4, 9]. Behaviorally, these animals show a general sluggishness, but interestingly, they exhibit an above normal vigor for dopamine-dependent actions like eating or drinking [21]. Tonic release of dopamine is believed to be the primary source of extra-cellular dopamine, and hence, a drop in the extra-cellular dopamine level suggests some kind of deficiency in the tonic dopamine generation. The exact reason of this deficiency in case of chronic hunger is still unknown.

Studies in humans and animals have revealed that the dopamine system is controlled by a powerful homeostatic mechanism. This mechanism compensates for

any changes in the level of extra-cellular dopamine by changing the rate at which tonic dopamine is synthesized and by changing the synaptic responsivity of the dopamine receptor neurons [4, 9, 12]. Once again, the exact mechanism of this restoration process is not completely clear. However, it has been observed that a period of sustained dopamine decrease is followed by the development of new synapses in the dopamine receptor neurons. This increase has been reliably replicated in a number of studies where the dopamine levels were chemically suppressed [12]. Recently, it has also been observed in studies where the dopamine levels were brought down by natural causes like food deprivation [9]. It is believed that receptor responsivity could increase even without any visible development of new synapses [9, 12, 21].

Thus, two opposing forces are at work in chronically hungry animals. The first mechanism decreases the dopamine level and the second mechanism compensates for the decreased dopamine levels. A side effect of the compensation mechanism is that the responsivity of the dopamine receptor neurons increases not only for the tonic dopamine, but also for the phasic dopamine release. Hence, over a period of sustained hunger, the animal's phasic dopamine system would become more and more responsive. This mechanism helps in explaining the selective increase in vigor for dopamine-dependent behaviors. The top part of Figure 1 summarizes these findings.

Next, we explore how hunger affects the action rate within an experimental session. It is observed that food consumption correlates with an increase in the dopamine level. This increase is seen only if the animal is hungry. Animals that have been fed ad-libitum do not show an increase in the dopamine level when they are fed [3]. Second, changing the dopamine level artificially does not change the animal's appetite – the total amount of food consumed by the animal remains the same, despite the changes in the dopamine level [25]. This evidence, when pieced together, suggests that the process of food consumption, along with the current hunger level, together act to dynamically modulate the extracellular dopamine level [32].

This dynamic regulation of the dopamine level seems to govern the action rate, an increase in the dopamine level causes an increase in the action rate and a decrease in the dopamine level causes a decrease in the action rate (Figure 1 bottom). First, there is evidence that the dopamine level directly correlates with the action rate [32]. Second, the animals that are administered with dopamine antagonists show slower rate of action [25]. Similarly, experiments in which the animals' dopamine level is enhanced show a more vigorous rate of action [8]. This evidence points towards the existence of a causal relation between the dopamine level and the action rate.

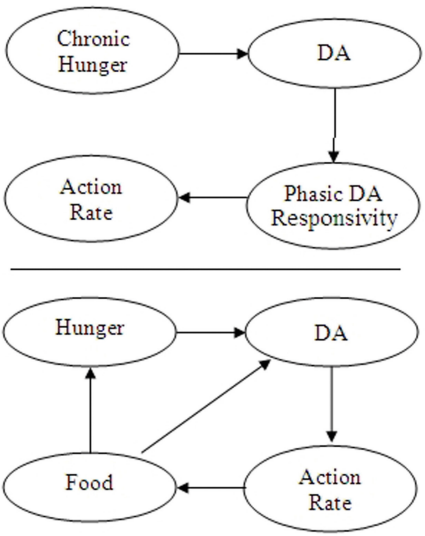


Figure 1. Top: Chronic hunger results in a decrease in the amount of extra-cellular dopamine. This triggers the homeostatic mechanism, which increases the responsiveness of the phasic dopamine system. This results in a more vigorous dopamine-related response. Bottom: Action rate is directly proportional to the dopamine level. The rate at which food is received depends on the action rate. Food consumpt ion decreases the the level of hunger. Hunger and food consumption affect the dopamine level.

2.1 Computational Models for Tonic and Phasic Activities of Dopamine

Numerous computational models have been proposed that describe the phasic activities of dopamine cells for learning and also show how this signal is used to maximize reward [29]. Variations in the concentration levels of dopamine affect a wide variety of behavior, some of which do not have any relation with phasic spiking activity of dopamine. Hence, for studying the influence of dopaminergic manipulation over response vigor, a model should incorporate the tonic level of dopamine also.

A mathematical model presented by Niv [19] demonstrates free-operant behavior. The model calculates reward based on response time and response rate in accordance with different reinforcement learning scenarios. Average rate of reward provides significant motivation and acts as an opportunity cost. The reward is dependent on the selected action, and the latency for performing that action. The goal of the subject is to select the best action-latency pair to maximize the reward in long run and also minimize the incurred cost per unit time [18].

2.2 The Leabra Framework

LEABRA (Local, Error-driven, Associative, Biologically Realistic Algorithm) [20] is a collection of computational formalisms for developing cognitive models that make contact with both observable behavior and detailed biological mechanisms. LEABRA models are constrained by our knowledge of processes at the level of membrane channels and individual neural functioning and also by our knowledge of gross brain anatomy and the role of various neurotransmitter systems. Synaptic weight learning in LEABRA includes a Hebbian learning algorithm, allowing for self-organization learning, and an error-correction learning algorithm formally related to the backpropagation of error technique. Error driven learning is performed using GeneRec, which is a generalization of the Recirculation algorithm, and approximates Almeida–Pineda recurrent backprop. LEABRA networks can also make use of a reinforcement learning algorithm based on the role of the dopamine neurotransmitter system in learning. By bringing all of these mechanisms together, LEABRA provides a single focal framework which supports a wide variety of connectionist concepts.

2.3 The Role of Dopamine

The striatum consists of two main types of cells with varying responses to phasic changes in dopamine (DA) that occur during error feedback. This causes two groups of striatal cells to independently learn positive and negative reinforcement values of responses, and ultimately acts to facilitate or suppress the execution of commands in the frontal cortex. The DA signal has to fluctuate substantially from its baseline levels in order to differentiate between response outcomes. Positive feedback results in an increase in tonic DA [27] whereas negative feedback lowers tonic DA levels [15]. These positive and negative fluctuations in levels of extracellular DA during feedback play a vital role in learning. Stimulation of the two types of DA receptors D2, and D1 have opposing (inhibitory and excitatory respectively) effects on long term potentiation (LTP). Expectedly, LTP is blocked by D1 antagonists and enhanced by D2 antagonists. Behavioral consequences of D1 antagonist administration include disrupted learning in appetitive conditioning tasks, whereas D2 antagonists promoted learning in the same. Hebbian or associative learning may be enhanced in the presence of DA, as this type of learning depends on the levels of activity of the cells in question. Thus, the efficacy of recently active synapses may be reinforced by a burst of DA acting as a teaching signal, leading to the learning of rewarding behaviors.

3 Mathematical Model

We first present an abstract model of animal conditioning that is based on the behavioral and neuroscientific results discussed in subsequent sections. While abstract mathematical models do suffer from the limitations of being somewhat removed from biology, they nevertheless serve as an excellent starting point and provide a less cluttered overview of the dynamics of the system. We acknowledge these limitations of our mathematical model and attempt to address these limitations in the subsequent sections with our neurocomputational model which is more firmly rooted in biology.

Our model consists of the following variables: Hunger represents the animal's hunger as a numerical score. EDA is a numerical representation of the extracellular dopamine level. ActionFreq signifies the action frequency. ChronicHunger signifies the chronic hunger, expressed as a percentage drop in the body weight. Reward is a binary value signifying whether the animal received a reward (a food pellet) or not. DAR is a numerical value representing the responsivity of the phasic dopamine system. In our model, time is divided into discrete steps, and the value of these variables is updated at each time step[1].

The action frequency in our model is directly proportional to the extra-cellular dopamine level as well as the responsivity of the phasic dopamine system. Hence, the action rate for the tth time step is computed as

$$\mathsf{ActionFreq}(t) = \alpha \times \mathsf{EDA}(t) \times \mathsf{DAR}, \tag{1}$$

where α is a model parameter. In the above expression, DAR does not depend on t. DAR changes over a time scale ranging from a few days to a few weeks. On the other hand, t, which represents the time period within a single experimental session, is too small to have any significant changes in the value of DAR.

In the above expression, the action rate is not shown to depend on the phasic dopamine activity. Neuroscientific studies show that the phasic activation of dopamine neurons is responsible for triggering the reward-related action in animals [28]. Phasic signals can be assumed to be incorporated implicitly since they are always required for action initiation. Also, the strength of these phasic signals, which varies during conditioning and extinction training, can be modeled by varying the value of the parameter α in the above expression.

The level of hunger during a training session decreases as a function of the cumulative reward (food) value:

$$\mathsf{Hunger}(t) = \mathsf{Hunger}(t-1) - \beta \times \mathsf{Reward}(t), \tag{2}$$

[1] The code for the mathematical model may be downloaded from `https://sites.google.com/site/lovekeshhome/Home/code.zip?attredirects=0&d=1theinternet`.

where β is a model parameter. According to the above expression, the level of hunger can either decrease or remain the same during an experimental session. The training sessions in the behavioral experiments typically last between 30 to 60 minutes. This time period is too short for the animals to digest the consumed food and become hungry once again. Also note that the food consumed during previous training trials does not affect the hunger during the next training trial, as successive sessions of behavioral experiments are conducted after sufficient time gap to eliminate that possibility.

As described in the previous section, food consumption leads to the increase of dopamine and this release is contingent on the animal being hungry. This relationship between food consumption, hunger and the level of dopamine is captured as follows:

$$\mathsf{EDA}(t) = \gamma \times \mathsf{EDA}(t-1) + \lambda \times \mathrm{Sig}(\phi \times \mathsf{Hunger}(t) \times \mathsf{Reward}(t)), \qquad (3)$$

where γ and λ are model parameters and $\mathrm{Sig}(\cdot)$ represents the sigmoidal activation function. In the above expression, $\mathsf{Hunger}(t) \times \mathsf{Reward}(t)$ will be zero at time steps when no reward is delivered. For time steps when reward is delivered, the value of the product will be equal to the level of Hunger. Hence, as per the above expression, the value of EDA will increase when a hungry animal gets a reward. Dopamine responsivity in our model changes as follows:

$$\mathsf{DAR}(T) = \mathsf{DAR}(T-1) + \theta \times \mathsf{ChronicHunger}(T), \qquad (4)$$

where θ is a model parameter. In the above expression, the use of T instead of t signifies the difference in time scales. t is used to represent time steps within an experimental session. In comparison, T denotes the number of days, as the change in dopamine responsivity does not happen over the period of a single experimental session. It requires a sustained food deprivation for a period of a few days to a few weeks.

We model the effects of dopamine depletion by scaling down the value of EDA. Hence, dopamine depletion to 80% of baseline in our model would mean that the value of EDA is 80% of the value that would be seen under normal circumstances.

Model parameters were determined through a combination of grid search as well as trial and error. The following values were used for the simulation results reported in this paper: $\alpha = 0.3$, $\beta = 0.01$, $\gamma = 0.985$, $\lambda = 0.22$, $\phi = 1.5$ and $\theta = 0.4$. The values of the variables were initialized as follows: $\mathsf{Hunger}(0)$ was set to 5.5 for hungry rats and 3 for pre-fed rats. $\mathsf{EDA}(0)$ was set to 0.5, and was restricted to range between 0.5 to 3.0. DAR was set to 1. As mentioned earlier, the value of Reward was either 1 (signifying the delivery of a food pellet) or 0 (no food) at each time step.

In the next two sections we discuss a related behavioral phenomenon and some simulations showing a qualitative match with the behavioral data.

3.1 Chronic Starvation

Behavioral studies show that the degree and duration for which the animals are starved has a direct correlation with the responses vigor. For example, Cagniard et al. [8] compared the responses of mice with varying degrees of starvation. Three sets of mice were used in their study, having 15%, 8% and 0% below baseline body weights. As expected, most vigorous response was seen in the mice maintained at 15% reduced body weight, followed by the mice maintained at 8% reduced weight followed by the normal mice. It is important to note that the mice learned the association between lever presses and food delivery before they were put on different food deprived regimes. Hence, the difference in the response rates cannot be attributed to the differences in the association strengths formed during conditioning. If these differences are not due to different association strengths, what could be the possible explanation?

Our model provides a simple explanation for these results. In our model, the effect of chronic hunger is captured by changing the dopamine responsivity (DAR) according to Equation (4). Since DAR is directly proportional to the action frequency, an increase in DAR results in a more vigorous response. Maintaining the animals at 10% below body weight for 5 days in our model resulted in an increase in the action rate from 34 per minute to 45 per minute.

Niv et al. [18] have proposed an abstract reinforcement learning model of animal conditioning. Their model incorporates the effects of chronic starvation by changing the utility of the reward. For example, a reward that is worth 10 units to some animal would be worth 15 units to a hungrier animal. The action vigor in their model is directly proportional to the reward utility. Hence, changing the utility leads to a change in the action vigor. It is important to note that in their model, the utility of a particular type of reward would result in increased vigor only for the actions that are associated with that type of reward. In other words, for example, a starved animal should not drink water more vigorously. This is different from our model, where a change in DAR predicts that the animal will act more vigorously, not only for food, but for all dopamine-dependent actions.

Niv et al. [18] do consider the evidence for a non-specific increase in action vigor in hungrier animals. They do so in the context of experiments where hungry animals are tested in a setup where they are free to perform some food-related action (like lever-press) as well as a drinking action. In such a setup, they explain away the increased vigor observed for the drinking action as the animal's desire to quickly revert to the food-related action. Our model predicts that the increased

vigor of drinking is not contingent upon the availability of any food-related action. Instead, it is due to the increased dopamine responsivity. Niv et al. [18] proposed a novel and promising theory to explain free-operant behavior. Our work tries to come up with a biologically plausible neurocomputational model to account for a similar phenomenon that digs deeper into the possible underlying processes.

3.2 Spontaneous Recovery

Spontaneous recovery is considered to be one of the most fascinating phenomena related to the extinction of conditioning. Here, the response vigor is seen to decrease during the extinction training and this decrement in the response vigor is seen to dissipate with the passage of time. This observed return of a portion of the originally learned behavior has been widely interpreted as evidence that extinction does not reverse the originally learned association [13, 23]. Many theories have been proposed as an explanation for the phenomenon of spontaneous recovery. One of the oldest explanations has been that the extinction-related associations get weakened with the passage of time due to the effects of non-specific interference. As another possibility, Bouton et al. [7] suggest that acquisition and extinction trainings are conducted in different temporal contexts and hence, these effects tend to average out as the animal moves into a "new" context with the passage of time. In a similar theory, Devenport [10] proposes that the response rate depends on the temporally weighted average reward value, with a higher weight given to the more recent events. Hence, just after extinction training, the weighted average is low, resulting in a less vigorous response. However, as the time passes, the temporally weighted average increases once again, resulting in increased responding.

In a typical spontaneous recovery experiment, the animals first undergo conditioning where an action is encouraged by associating it with a reward. This is followed by extinction training, where the action rate is seen to drop due to the withdrawal of the reward. Finally, after a period of rest, the animal is once again tested for its response rate. It is typically seen that the passage of time results in an increase in the response rate. Rescorla [22] incorporated an additional stage in the above described experiment. After the extinction training, he subjects the animals to a period of reacquisition. This was followed by the rest period and the test for response rate. He found that the animals responded with rates greater than those seen at the end of the reacquisition phase. He attributed this to the phenomenon to spontaneous recovery. He used this result as an evidence in support of the dual pathway theory of animal conditioning. He suggested that extinction could involve the formation of a separate decremental association, and not a reversal of acquisition related association. Similarly, reacquisition might involve a slight strengthening of the acquisition related association and possibly a slight

weakening of the extinction related association. Most of the extinction related association, however, must have survived through the reacquisition training. Over the rest period, this extinction related association must be going through a phase of decay or interference, causing the spontaneous recovery. Rescorla [23] wondered if there is a scenario where an opposite effect would be observed, i.e. the response rate decreases with the passage of time. He conducted a series of experiments, manipulating the order and duration of conditioning, extinction as well as rest periods. In all the cases, he found that the rate of responding after the rest period was greater than the rate of responding before. In fact, in some of his experiments, after the period of rest, the animals responded at rates that were higher than the rates ever reached during any of the training sessions. From this, Rescorla concluded that there must be something peculiar about the extinction related association that makes it weaker with the passage of time.

It should be noted that throughout his experiments, Rescorla maintained the animals at 80% of this normal body weight. This was true even for the rest period before the test of spontaneous recovery. The period of rest in his experiments was five days. Hence, it is likely that the phasic dopamine responsivity of these food deprived animals slowly increased during the rest period, causing a more vigorous responding during the spontaneous recovery test.

As explained before, in our model, the effects of extinction are captured by decreasing the value of the parameter α in Equation (1). Decreasing α from the default value of 0.3 to 0.15 results in a decrease in the lever press rate from 34 per minute to 7 per minute. Now, if we maintain the model 20% below the baseline body weight for a duration of 5 days, the lever press rate increases to a value of 16 per minute.

It should be noted that this explanation does not eliminate the need for the other theories of spontaneous recovery. In Rescorla's experiments, extinction training was conducted by the omission of rewards. Other conditioning studies, in which the extinction training was conducted by punishing the animals with a foot shock or some other undesirable event, report a different behavior – they report the spontaneous recovery of the fear response. Evidence suggests that fear responding is not governed by dopamine firing and we do not yet have a clear understanding of the mechanisms that might underlie the formation of fear related associations. In yet another variant, animals are trained to press a lever for food delivery. Extinction training involves the omission of reward for the lever pressing behavior in addition to reward delivery for a previously unrewarded behavior (like the pressing of some other lever). Over time, the animals stop pressing the first lever and start pressing the second lever. After a period of rest, it is seen that the behavior of pressing the first lever returns, such that animals now start choosing both the levers almost the same number of times. While increase in phasic responsivity is

still possible in such experiments, additional phenomena like memory consolidation [31] may be playing some role as well.

Finally, a dopamine based explanation for the phenomenon of spontaneous recovery mitigates its applicability as an evidence for the dual-association hypothesis. Other evidence, however, still remains best explained via the dual association hypothesis [13] and hence, the dual-hypothesis cannot be completely ruled out.

4 Neurocomputational Model

Frank [11] developed a relevant model to simulate the dynamic behavior of dopamine in the basal ganglia (BG). This model is closely related with our work. Frank simulated the gating based motor response selection mechanism by modeling D1 and D2 receptors in the striatum. He further simulated the two main pathways, namely the direct and indirect pathway via which striatal neurons project to the Globus Pallidus and Substantia Nigra. The direct pathway facilitates the execution of responses, whereas the indirect pathway inhibits them [11]. See Figure 2 for a pictorial description of this circuitry.

The neurocomputational model of the basal ganglia proposed by Frank does not account for the level of hunger experienced and the resulting interactions. Franks model was therefore modified to incorporate the concept of the food reward mechanism.

4.1 The Function of Acetylcholine

Acetylcholine neurons comprise 1–2% of the total neurons in the striatum. They are large, aspiny interneurons which project to medium spiny neurons. The topography of the projections is similar to that of doapminergic neurons from the substantia nigra [16]. The direct pathway is made up of substance P containing neurons whereas neurons in the indirect pathway contain enkephalin. The substance P type neurons express the D1 receptor as well as the muscarinic M1 and M4 receptors whereas the enkephalinic neurons mainly express the D2 and M1 receptors, with only 39% having the M4 receptor [6]. Recent evidence suggests that the action of ACh in the striatum is opposite to that of DA in that it stimulates the indirect pathway and suppresses the direct pathway.

Fos, the product of the proto-oncogene *c-fos*, is a marker of neuronal activity [24]. The study of Fos immunoreactivity has shown the differential action of DA and ACh on the two different pathways in the striatum. DA selectively facilitates the activity of the substance P type striatonigral neurons via the D1 receptors while suppressing the activity of enkephalinic striatopallidal neurons via D2 eceptors. At

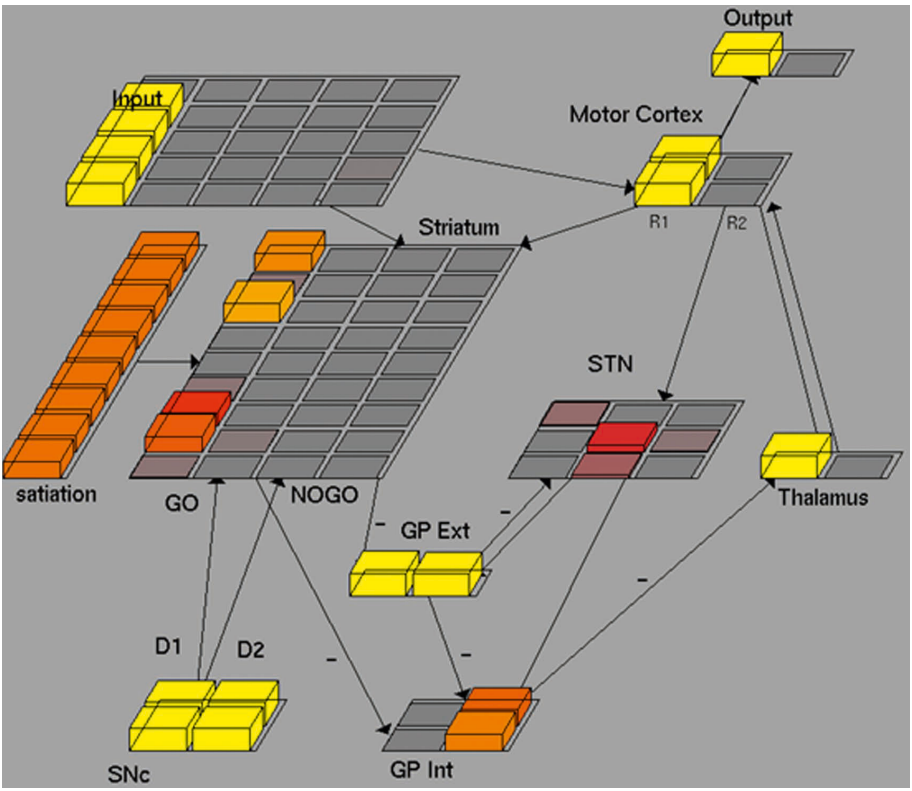


Figure 2. Neurocomputational model showing the satiation layer and connection between satiation layer and striatum.

the same time, muscarinic receptors stimulates the enkephalinic neurons independently of dopamine while suppressing the substance P type neurons [5].

Thus, ACh and DA maintain a dynamic balance in the BG circuitry, including the NAc which is also involved in feeding behavior. Microdialysis studies on free feeding rats show that the level of extracellular acetylcholine in the NAc peaks just after the maximum food intake [17]. Further studies conducted on rats fitted with gastric fistulas demonstrated that the peaking of extracellular acetylcholine is not observed in rats whose stomach contents were drained during the meal. The rats with the fistulas opened also ate more than the ones whose fistulas were closed. The rise in the level of extracellular DA was observed in both the groups [2].

In the light of the above evidence, it is proposed that the rise in tonic ACh serves as a marker for satiety. The peaking of extracellular ACh in the NAc is the starting point of the cessation of feeding.

4.2 Satiation Layer

A satiation layer is added to the model to specify the level of satiation during the food intake. The unit weights of this layer are modified via by scripts. Whenever the model chose the correct response, the unit weights of satiation layer underwent an increment (corresponding to a decrease in hunger). The satiation layer represents the extracellular ACh which opposes the action of DA by suppressing the GO pathway and exciting the NO-GO pathway with the overall effect of stopping the system from responding to food. Just as the levels of tonic ACh increases with the consumption of food [17], the satiation layer activity is made to increase with every correct response of the model, and decrease with every incorrect response. The rates of increase and decrease are kept asymmetric to allow for stable model dynamics.

5 Simulations with the Neurocomputational Model

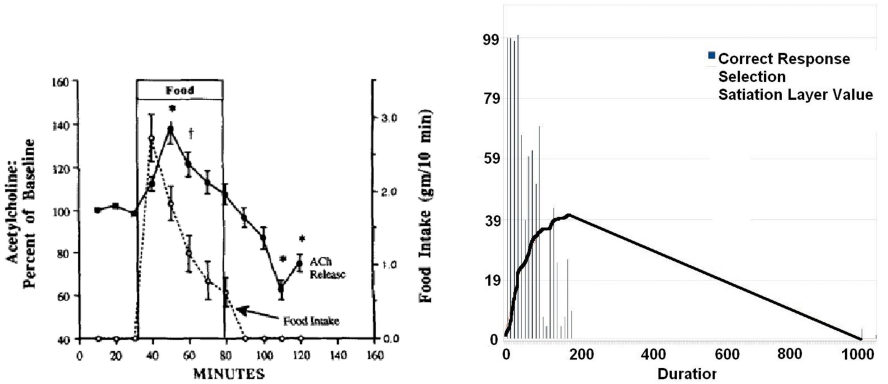
5.1 Correlation Between ACh and Feeding

Experiments conducted by Mark et al. [17] on rats demonstrated a peaking of extracellular ACh in the NAc soon after the maximal rate of food intake. The samples were collected in 10 minute intervals by microdialysis. The animals had been food deprived before the experiments (Figure 3 (a)). Feeding caused a 38% increase in extracellular ACh in the NAC. Dopamine was also increased in the NAC (48%).

The model intended to simulate the rise in the acetylcholine level due to food intake. First, the model was trained to respond correctly by running the simulation with very low values of satiation. In the simulation, the effect of tonic ACh was implemented by incrementing the satiation layer value by 0.05 units for each correct response of the model. The rate of decrement was set to 0.005 units for each incorrect response. The asymmetric response was necessary for a stable model dynamics (Figure 3 (b)). The simulation results correctly depict the rise in the ACh levels soon after the maximal rate of food intake followed by cessation of feeding. This is followed by the decrease in ACh levels to baseline levels over a period of time during which no feeding takes place.

5.2 The Effect of Draining the Stomach Contents

A set of novel experiments on rats fitted with gastric fistulas gives the strongest evidence of the role of ACh in the signalling of satiation [17]. Two sets of rats fitted with gastric fistulas were kept on an alternating 12 hour feeding and starvation schedule. In the first hour of the feeding schedule, the rats were given free access

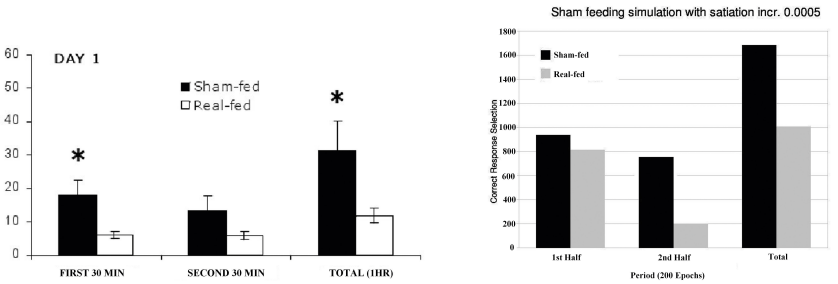


(a) (b)

Figure 3. (Experimental) (a) Extracellular levels of ACh in the NAC before, during, and after five 10-min intervals of food access. Subjects ($n = 10$) were deprived of food for 20 h before food presentation. ACh peaked in the 10-min interval immediately following maximal food intake ($*p < 0.01$; $\dagger p < 0.05$). ACh levels are expressed as the percent ($\pm$SEM) of the mean of three baseline samples [17]. (Simulation) (b) Graph showing the relation between feeding and the satiation layer value. The bars represent the number of correct responses in a session of 100 total responses. The line represents the satiation layer level in arbitrary units.

to sucrose solution. During this time, one set of rats had their stomach contents drained by keeping the fistulas open, while the other set was fed in the normal manner with the fistulas closed. It was observed that the rats having their stomach contents drained did not have the peaking of extracellular ACh in their NAc, compared to the ones which had the fistulas closed. It was observed that there was a significant difference between the amount of food consumed by the two groups in that duration. The rats with the fistulas open consumed greater amount of food than the ones which had them closed (Figure 4 (a)).

The neurocomputational model was run for both the cases. The model was trained, the initial satiation level set to 0 units and the simulation started. Simulations were run up to 450 epochs and the data is analyzed. The rate of decrement of satiation after each incorrect response is kept constant at 0.005 units. For the rats with closed fistulas, the rate of increment was kept at 0.05 units. For the rats with the fistulas open, results were obtained with the rate of increase set at 0.0005 units (a ratio of $1/100$ w.r.t. to the "real-fed" case). The "sham-fed" group consumed a significantly higher amount of food (Figure 4 (b)).



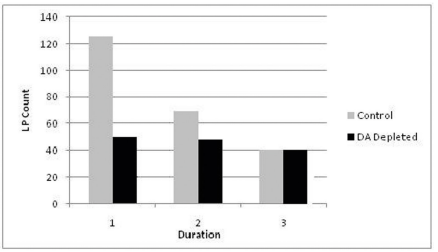
(a) (b)

Figure 4. (Simulation) (a) Sucrose intake measurements (in ml) during the first hour of feeding of sham-feeding and real-feeding rats with access to sugar. The asterisks indicate differences between groups ($P > 0.05$) [17].
(Simulation) (b) The figure showing the correct response selection in experiments with a satiation layer increment value of 0.0005 for the sham-fed vs. 0.05 for the real-fed.

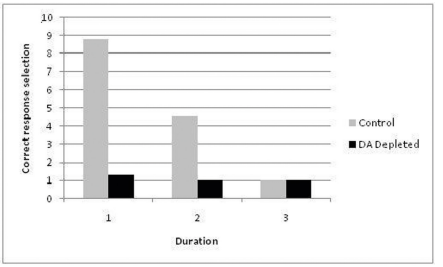
5.3 Changes in Action Frequency Within a Session

Salamone et al. [26] conducted a number of conditioning experiments to explore the relation between dopamine level and response vigor. In one set of experiments, using a continuous reinforcement schedule, they measured the number of lever presses over successive 15-minutes intervals. They found that the maximum number of lever presses occurred during the first 15-minutes. The number of lever presses decreased over the next two 15-minutes intervals as shown in Figure 5 (a). They conducted the same experiment with rats that were administered with dopamine antagonists. The number of lever presses over the first 15-minutes period was significantly smaller in dopamine depleted rats as compared to the control rats. However, the number of lever presses over the next two 15-minutes intervals was comparable to that of the control animals.

The neurocomputational model explained in Section 2 is executed for 620 epochs. After 170 epochs, the network learns for the appropriate output and the network error goes to zero. Subsequently, three consecutive sessions of 150 epochs each are run with a constant increase in the unit weights of satiation layer. It is found that the number of correct response selections decreases in successive sessions. This is because the hunger decreases with the increase in weights of the satiation layer units which directly affect the GO units of striatum layer. In case of dopamine depletion, the number of correct response selection is much smaller in comparison to the control case for the initial phase of the simulation. After the sec-



(a)



(b)

Figure 5. Changes in action frequency. (a) Behavioral results. (b) Neurocomputational model results.

ond and third phases, the number of correct responses selected further decreases. The results for control and dopamine depleted case are compared in Figure 5 (b).

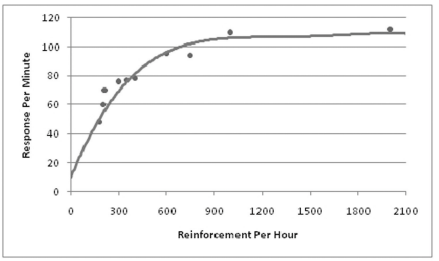
5.4 Effect of Reward Rate Manipulations

Herrnstein [14] reports the relation between the rate of reward and the rate of action in a random interval schedule of reinforcement. In this schedule, the first action after a specific time interval is rewarded. In such experiments, it is observed that the action rate increases with the reward rate, asymptotic to a high value (Figure 6 (a)).

The simulations run using the neurocomputational model described in Section 4. The results are shown in Figure 6 (b). The model uses an initial satiation layer activity value of zero. Simulations are executed for rewards after every 1, 2, 3, 4, 5 and 6 epochs. As shown in Figure 6 (b), for lower reward rates the model gets rewarded for each correct response and gets sated quickly. For higher reward rates, a proportion of the correct responses would not yield reward and hence the network needs a large number of correct responses to achieve satiation. Due to the slow learning rate, the network selects a large number of correct responses.

5.5 Effect of Reward Ratio Manipulation in Case of Pre-Fed Rats

The term reward ratio in appetitive conditioning is defined as the number of lever presses the subject is required to perform in order to obtain a food pellet. Aberman et al. [1] conducted an experiment to measure the variation in the response rate with changing reward ratios for pre-fed rats. As shown in Figure 7 (a), both control and pre-fed rats demonstrate increasing response rate with increasing reward ratio. The response rate plateaus for higher reward ratios. One reason for this could be that there is a physical limit to the number of lever presses in one experimental



(a)

(b)

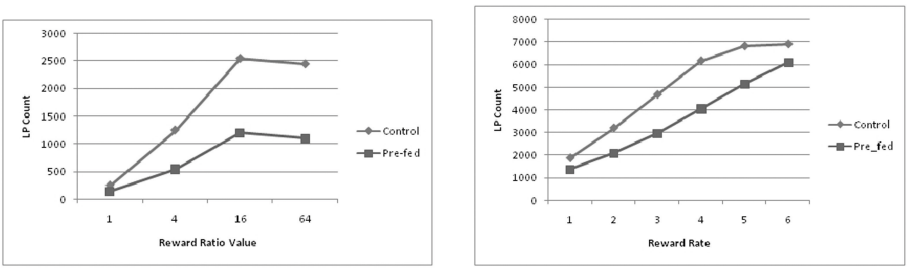
Figure 6. Experimental results showing an increasing trend of response rate for increasing the reinforcement rate. (a) Behavioral results. (b) Neurocomputational model results.

session. Also, for higher reward ratios, the curve is declining because for higher reward rate, the rats become less responsive. They might be satiated or they may not learn correct actions due to less frequent reward for pressing the correct lever. The pre-fed rats show consistently less active lever presses for all reward ratios greater than 1. For the pre-fed rats (i.e. sated rats), the number of lever presses is less in comparison to control rats.

To simulate these experiments the model is initially trained for the correct response. For reward ratio of 2, the script was modified so that the network gets one reward for each alternate phasic burst in SNc layer. A similar procedure is followed for reward ratios 3, 4, 5 and 6. For pre-fed experiments, we chose a higher initial activation value for satiation layer. The results match the behavioral experiments as shown in Figure 7 (b). For a higher reward ratio, the model exhibits a higher action rate.

6 Conclusions and Future Work

This paper explores the role of hunger in the learning of dopamine dependent behaviors, and presents both an abstract mathematical model and a more biologically plausible neurocomputational model that incorporates the role of dopamine and acetylcholine in the acquisition of hunger related associations. Several challenges remain to be incorporated into our model. Firstly, our model is based solely on the evidence for correlations between hunger, dopamine levels and the response rates. The mechanism underlying these correlations remains to be explored. Second, we have only explored the effects of hunger on a previously learned behavior. The possible role of hunger and dopamine levels in the learning process also needs to



(a) (b)

Figure 7. Experimental results showing an increasing trend of response rate for increasing the reinforcement ratio. (a) Behavioral results. (b) Neurocomputational model results.

be explored. Further, several factors other than hunger could possibly affect the dopamine levels. For example, there is evidence that stress results in the release of tonic dopamine. Also, action frequency could be affected by several factors other than just the the dopamine level. Fatigue, other possible rewarding actions, other available rewards, time gap between action and reward and the amount of effort required for performing the action are some of the possible factors. Finally, dopamine could be responsible for more than just governing the action rate. It could govern the action timing, action selection, reward selection, perception, and motor execution. Our future models aim to incorporate these factors after thorough experimentation.

Bibliography

[1] J. E. Aberman and J. D. Salamone, Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement, *Neuroscience* **92** (1999), 545–552.

[2] N. M. Avena, P. Rada, N. Moise and B. G. Hoebel, Sucrose sham feeding on a binge schedule releases accumbens dopamine repeatedly and eliminates the acetylcholine satiety response, *Neuroscience* **139** (2006), 813–820.

[3] V. Bassero and G. Di Chiara, Differential influence of associative and nonassociative learning mechanisms on the responsiveness of prefrontal and accumbal dopamine transmission to food stimuli in rats Fed ad libitum, *Journal of Neuroscience* **17** (1997), 851–861.

[4] N. T. Bello, K. L. Sweigart, J. M. Lakoski, R. Norgren and A. Hajnal, Restricted feeding with scheduled sucrose access results in an upregulation of the rat dopamine transporter, *Americal Journal of Physiology – Regulatory, Integrative and Comparative Physiology* **284** (2003), 1260–1268.

[5] V. Bernard, B. Dumartin, E. Lamy and B. Bloch, Fos immunoreactivity after stimulation or inhibition of muscarinic receptors indicates anatomical specificity for cholinergic control of striatal efferent neurons and cortical neurons in the rat, *European Journal of Neuroscience* **5** (1993), 1218–1218.

[6] V. Bernard, E. Normand and B. Bloch, Phenotypical characterization of the rat striatal neurons expressing muscarinic receptor genes, *Journal of Neuroscience* **12** (1992), 3591–3600.

[7] M. E. Bouton, Context and behavioral processes in extinction, *Learning and Memory* **11** (2004), 485–494.

[8] B. Cagniard, P. D. Balsam, D. Brunner and X. Zhuang, Mice with chronically elevated dopamine exhibit enhanced motivation, but not learning for food, *Neuropsychopharmacology* **31** (2006), 1362–1370.

[9] K. D. Carr, Y. Tsimberg, Y. Berman and N. Yamamoto, Evidence of increased dopamine receptio signaling in food-restricted rats, *Neuroscience* **119** (2003), 1157–1167.

[10] L. D. Devenport, Spontaneous recovery without interference: Why remembering is adaptive, *Animal Learning and Behavior* **26** (1998), 172–181.

[11] M. J. Frank, Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism, *Journal of Cognitive Neuroscience*, **17** (2005), 51–72.

[12] A. A. Grace, Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia, *Neuroscience* **41** (1991), 1–24.

[13] A. Gupta and D. C. Noelle, Lateral inhibition explains savings in conditioning and extinction, in: *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (2005), 309–314.

[14] R. J. Herrnstein, On the law of effect, *Journal of Experimental Analysis of Behavior* **13** (1970), 243–266.

[15] J. R. Hollerman and W. Schultz, Dopamine neurons report an error in the temporal prediction of reward during learning, *Nature Neurosci* **1** (1998), 304–309.

[16] P. N. Izzo and J. P. Bolam, Cholinergic synaptic input to different parts of spiny striatonigral neurons in the rat, *The Journal of Comparative Neurology* **269** (1988), 219–234.

[17] G. P. Mark, P. Rada, E. Pothos and B. G. Hoebel, Effects of feeding and drinking on acetylcholine release in the nucleus accumbens, striatum, and hippocampus of freely behaving rats, *Journal of Neurochemistry* **58** (1992), 2269–2274.

[18] Y. Niv, N. D. Daw and P. Dayan, How fast to work: Response vigor, motivation and tonic dopamine, in: *Proceedings of Neural Information Processing Systems*, MIT Press (2005), 1019–1026.

[19] Y. Niv, N. D. Daw, D. Joel and P. Dayan, Tonic dopamine: opportunity costs and the control of response vigor, *Psychopharmacology* **191** (2007), 507–520.

[20] R. C. O'Reilly and Y. Munakata, *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*, The MIT Press, 2000.

[21] E. N. Pothos, I. Creese and B. G. Hoebel, Restricted eating with weight loss selectively decreases extracellular dopamine in nucleus accumbens and alters dopamine response to amphetamine, morphine and food intake, *Journal of Neuroscience* **15** (1995), 6640–6650.

[22] R. A. Rescorla, Retraining of extinguished pavlovian stimuli, *Journal of Experimental Psychology: Animal Behavior Processes* **27** (2001), 115–124.

[23] R. A. Rescorla, Spontaneous recovery of excitation but not inhibition, *Journal of Experimental Psychology: Animal Behavior Processes* **31** (2005), 277–288.

[24] S. M. Sagar, F. R. Sharp and T. Curran, Expression of c-fos protein in brain: metabolic mapping at the cellular level, *Science* **240** (1988), 1328–1331.

[25] J. D. Salamone and M. Correa, Motivational views of reinforcement: Implications for understanding the behavioral functions of nucleus accumbens dopamine, *Behavioral Brain Research* **137** (2002), 3–25.

[26] J. D. Salamone, P. Kurth, L. D. McCullough and J. D. Sokolowski, The effects of nucleus accumbens dopamine depletions on continuously reinforced operant responding: contrasts with the effects of extinction, *Pharmacology Biochemistry and Behavior* **50** (1995), 437–443.

[27] W. Schultz, P. Dayan and P. R. Montague, A neural substrate of prediction and reward, *Science* **275** (1997), 1593–1599.

[28] W. Schultz, Multiple reward signals in the brain, *Nature Reviews, Neuroscience* **1** (2000), 199–207.

[29] R. S. Sutton and A. G. Barto, Reinforcement learning: an introduction, *IEEE Transactions on Neural Networks* **9** (1998), 1054–1054.

[30] L. Vig, A. Gupta and A. Basu, On the relation between hunger dopamine and action rate, in: *Proceedings of the Indian International Conference on Artificial Intelligence*, Springer (2011), to appear.

[31] J. T. Wixted, The psychology and neuroscience of forgetting, *Annual Review of Psychology* **55** (2004), 235–269.

[32] A. M. J. Young, M. H. Joseph and J. A. Gray, Increased dopamine release in vivo in nucleus accumbens and caudate nucleus of he ratt during drinking: a microdialysis study, *Neuroscience* **48** (1992), 871–876.

Received June 30, 2011.

Author information

Lovekesh Vig, School of Computational and Integrative Sciences,
Jawaharlal Nehru University, New Delhi 110067, India.
E-mail: lovekeshvigin@gmail.com

Ashish Gupta, Department of Electrical Engineering and Computer Science,
Vanderbilt University, Box 1679 Station B, Nashville, TN 37235, USA.
E-mail: ashish2240@gmail.com

Abhinandan Basu, School of Computational and Integrative Sciences,
Jawaharlal Nehru University, New Delhi 110067, India.
E-mail: abhinandan.basu@gmail.com