Chunking of Action Sequences in the Cortex-Basal Ganglia System

Johannes Hjorth
f97-jhj@f.kth.se

TRITA-NA-Eyynn

Abstract

The cortex-basal ganglia system is important in the learning of new motor behaviour. Recent experimental studies suggest that the basal ganglia are mainly involved in the initial stages of learning. We propose that, while the basal ganglia are important for the acquisition of new behaviour, the repetition of the behaviour causes consolidation of the action sequence as a unit in the motor cortex. In this work we try to simulate this phenomenon and create a model of the cortex-basal ganglia system. This models the neural control of a rat running through a labyrinth. Initial results are promising and future versions of the prototype will allow comparison with recorded biological data from live rats.

Sammanbindning av handlingssekvenser i kortiko-basalgangliära systemet

Examensarbete

Sammanfattning

Det kortiko-basalgangliära systemet tros ha en viktig roll vid inlärning av nya motoriska sekvenser. Nya experiment visar på att de basala ganglierna är involverade främst i första stadierna av inlärning. Vår hypotes är att det är de basala ganglierna som lär sig motoriska sekvenser, men att de sammankopplade motorsekvenserna därefter lagras i kortex. I den här rapporten försöker vi simulera detta fenomen genom att skapa en prototyp av kortiko-basalgangliära systemet. Vår model simulerar en råtta som springer i en labyrint. De första resultaten ser lovande ut och simulerade data från framtida versioner av vår prototyp kommer att kunna jämföras med biologiska data från levande råttor.

Contents

Chapter 1

Introduction

This report is based on a novel hypothesis about how habitual behaviour is represented and learned in the brain. It originates from the work of Graybiel (1998) and parts of it have been published by Djurfeldt et al. (2001). The traditional picture is that the action is learned initially in cortex, and then later stored in and accessed through the basal ganglia, but we suggest that the basal ganglia learn a sequence needed through trial and error, but once the right sequence has been acquired it is consolidated in the cortex. The goal for this report is to present a model that shows that elements similar to neurons can perform such a task and that the learning-transfer process can be captured by the cortico-basal ganglia model.

We simulate the interaction between the cortex and the basal ganglia of a rat running through a grid world maze shaped like a T. We are here trying to replicate the experimental conditions in Jog et al. (1999). Our ultimate goal is to be able to compare simulated data from the model with data from real rats recorded by inserting multiple tetrodes in the striatum and monitoring the activity of populations of neurons over periods of days.

Below is a picture of a human skull painted by Leonardo da Vinci around 1500. Our understanding of the brain has gone a long way since then, but his immortal paintings and sketches are a beautiful reminder of the curiosity that drives us in our research.

Chapter 2

Neurobiological Background

Neural signals are transmitted from the axon of one neuron to the dendrite of another by chemical transmitters passing across the synapse, a small gap between the two neurons. Each excitatory input to the dendrites temporarily raises the potential of the neuron. If a threshold is passed the neuron fires, sending out a signal along the axon to other neurons. The changes of the potential in the cell can be described by the differential equations of a leaky integrator. We distinguish between two types of neurons: interneurons which act locally and projection neurons which have a long axon and send signals to other regions.

2.1   The Basal Ganglia

The basal ganglia are centrally placed in the brain (Figure 2.1). They play an important role not only in selection and execution of learned behaviour, but also in acquisition of new behaviours and maintenance of old ones (Wickens, 1997).

In figure 2.2 the different parts of the basal ganglia can be seen. The striatum, which is the main input centre of the basal ganglia, receives inputs from all major parts of the cortex. Two main pathways lead from the striatum – the direct and the indirect pathway, the former has an overall excitatory effect, on the corticothalamic network that activates motor sequences, while the latter inhibits them. Dopamine, a neurotransmitter often associated with learning, facilitates the direct pathway and inhibits the indirect. Imbalance between the two pathways can result in hyper- or hypo-kinetic disorders (Parent et al., 2001).

The level of dopamine is increased when an unexpected reward is given, it is at base levels if the anticipated reward was given, and it is lower than normal if the expected reward was not given. It is interesting to note that the depression when an anticipated reward was not given happens at the time the reward was expected (Schultz et al., 1997).
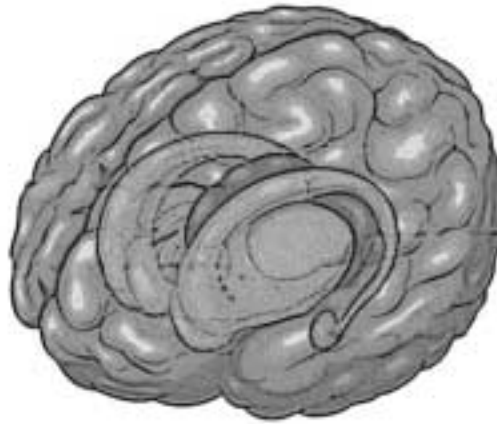
Figure 2.1. The Basal Ganglia from Neuropsychology by Marie T, Banich and Houghton Mittlin, 1997. The image has been retouched to remove text.

2.1.1 The Direct and Indirect Pathways

The direct pathway is a projection system that connects the striatum to the internal segment of the globus pallidus (GPi) and the Substantia Nigra (SNr) through inhibitory connections, see figure 2.2. These two regions are physically separated but have similar functionality, and these in turn inhibit the thalamus. The double inhibition means that the direct pathway has a disinhibitory effect on the thalamus, that is it reduces the inhibition acting on the thalamus. From the thalamus projection neurons send signals to lower level motor neurons and also connect back to cortex, completing the circuit.

The indirect pathway goes from the striatum to the external segment of globus pallidus (GPe), which in turn connects to the subthalamic nucleus (STN). Both these connections are inhibitory. From the STN strong excitatory connections go to both the GPi and the SNr, and they in turn inhibit the thalamus as mentioned above. This means that the indirect pathway has an overall inhibitory effect on the motor activity.

2.1.2 The Striatum

The striatum is dominated by inhibitory projection neurons, located in a single layer. From experiments it has been shown that the regular activity is relatively low. Individual neurons can be silent for a few seconds up to hours. These periods of inactivity are interrupted by episodes of moderate activity. There are neurons which have been found to fire in anticipation of reward and others fire prior to initiating learned behavior. There is also a level of abstraction, with a significant portion of the striatal neurons coding for direction of movement, rather than for details in muscle activity.
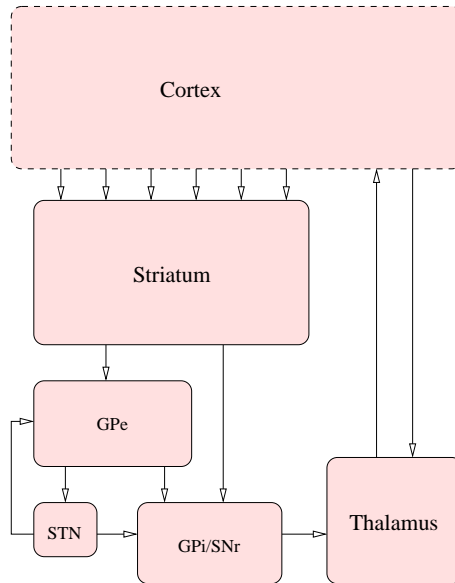
Figure 2.2. Schematic diagram of the basal ganglia. Abbreviations: GPe, external segment of globus pallidus; GPi, internal segment of globus pallidus; STN, subthalamic nucleus; SNr, substantia nigra, pars reticulata (Wickens, 1997).

From recordings of the striatum in live rats, gradual changes in the firing patterns of projection neurons have been found as the rats acquire new behaviours (Graybiel, 1998). Input from large areas of the cortex converges on striatal projection neurons, each making a small contribution, which suggests either a high rate of firing from a few neurons or coherent network activity from large number of neurons, is needed to make the projection neurons fire. The firing in the striatum is very dependent on specific contexts which implies that it corresponds to learned behaviour.

Different kinds of sensory input from the same regions of the body are usually projected onto the same area of the striatum.

In addition to the projection neurons there are also interneurons in the striatum, which are fewer than the projection neurons but still serve an important role since they can locally create relatively high concentrations of neuromodulators like acetylcholine.

The striatum also receives dopaminergic input from Substantia Nigra Compacta (SNc) and the ventral tegmental area. The dopamine axons branch extensively and terminate in the striatum close to where the input is received from the cortex. Here they are able to regulate the efficiency of the synapses.

2.1.3   The Pallidal-Subthalmic Complex

The neurons in GP have tonic activity with high frequency firing, interrupted by brief reductions in the firing rate, corresponding to the short bursts of inhibition

from the striatum.

The subthalamic nucleus (STN) mentioned earlier as part of the indirect pathway is sometimes labelled as the driving force of the basal ganglia because of its strong excitatory effect on GPi and SNr, the output regions of the basal ganglia, that in turn inhibit the thalamus. This constant inhibition of the thalamus keeps the motor system inactive, but ready to respond. When it is time to trigger a behaviour a wave of increased inhibition is seen. This is followed by the release of inhibition from the behaviour desired, while the other ones are kept inhibited. Afterwards the inhibition is once again strengthened for a period before it is released back to normal levels.

2.2   Neurological Diseases

Imbalance between the direct and indirect pathways can result in a changed activity in the GPi/SNr which could account for the hypo- and hyperkinetic disorders sometimes associated with the basal ganglia (Parent et al., 2001).

It has been suggested that Bradykinesia or akinesia that are characteristic of Parkinson's disease result from increased inhibition of thalamic premotor neurons due to excessive excitatory drive from the STN to GPi/SNr. This would result from increased inhibition of the GPe due to a loss of striatal dopamine leading to the disinhibition of the STN.

At the other end there is the hyperkinetic movements associated with Huntington's disease. Here the problem is believed to be the opposite, too weak inhibition from the striatum onto the GPe, due to degeneration of striatal projection neurons, leads to too high inhibition of STN, and a lack of excitation for GPi/SNr which results in increased motor activity.

Chapter 3

Simulation Models

The purpose of this chapter is to give a short presentation of the different components of our model and how they work.

Our hypothesis is that the basal ganglia learn the new behaviour and that once it has been learned it is transferred to the cortex where it is consolidated. We seek to investigate if it is possible to create a mechanism using neuron-like elements that can learn the required sequences in one network, then transfer control to a second network.

Below we will give a description of the various networks and the mechanisms that make this transfer of learned behaviour possible. We also survey a few models that are interesting to study for the further development of our model. Future versions of the cortex-basal ganglia model are likely to be spiking to allow better comparison with biological data.

3.1   Basal Ganglia Models

There is a plethora of different models that try to better mimic the function of the basal ganglia. The level of detail has increased over time. Houk et al. (1995) presented one of the first models. The problem with it was that it did not account for the timed depression of dopamine when an expected reward was omitted. This was solved in Suri and Schultz (1998, 1999) where the timing mechanism was implemented by representing each stimulus using a set of neurons, each of which was active for a different duration rather than using a single prolonged inhibition. Using their network they showed that relatively complex tasks could be solved, however the novelty responses, generalization responses and some temporal aspects in reward prediction was achieved by setting arbitrary values, such as initializing synaptic weights with specific values and using different learning rates for different synapses. Contreras-Vidal and Schultz (1999) provided a model closer to the anatomy of the basal ganglia. They also proposed two error signals, one for errors in predicting the timing of reward and another for type and amount of reward. Further they proposed that the fast excitatory response to conditioned stimuli and the time delayed

inhibition in response to unconditioned simuli are mediated by different pathways. Brown et al. (1999) continued the work but instead of assuming that the striosomal neurons generated a spectrum of timing signals in response to sensory input they used an intra-cellular calcium-dependent timing mechanism.

3.2 Actor Critic Model of the Basal Ganglia

In this section we are describing the basal ganglia model used in this work. The model is based on the actor-critic architecture (Sutton and Barto, 1998) which we will describe here.

The dopamine in the basal ganglia appears to act like a regulator of anticipated reward (Schultz et al., 1997). When an unexpected reward is given the dopamine levels increase, when a reward is given as expected the dopamine levels are normal, and when the expected reward did not happen the levels are below normal. In the initial stages the dopamine levels are directly in response to the reward, but after learning the cue triggers the internal reward mechanism.

This behaviour bears a close resemblance to the temporal difference learning (Sutton and Barto, 1998) which often uses an actor that follows a policy and a critic that evaluates the action. Here the temporal difference given by the critic would correspond to the output from the dopamine neurons in the VTA and SNc (Schultz et al., 1997). The temporal difference error for the action a_t taken in state s_t is evaluated as,

$$\delta_t \quad = \quad r_{t+1} + \gamma(V(s_{t+1}) - V(s_t)) \tag{3.1}$$

where r_{t+1} is the reward, γ is the discount-rate, usually to 1.0 or slightly below, and $V(s)$ is the value function at state s.

The general theme for models of the critic is the dopaminergic neurons activity that bears a close resemblance to temporal difference learning. There are however still questions regarding how to reproduce the dynamics of firing to rewards, reward predicting stimuli and novelty. Joel et al. (2002) propose the idea of using an evolutionary computational approach to find candidate architectures that maximize the utility of the critic under anatomical and functional constraints. Work has also been carried out on the actor side, which is based around the dopamine dependent long-term synaptic plasticity in the striatum.

The model we use consists of two separate structures, one structure called the actor which implements a policy and an estimated value function called the critic. The policy is a method to choose action based on the state, and the value function predicts how good it is to be in a certain state. The critic criticizes the actions taken by the actor by returning a temporal difference error in the form of a scalar. If this scalar is positive, the tendency to select the action should be strengthened and if negative it should be reduced.

The easiest way to choose the best action is by using the greedy policy that tries to maximize the reward at each step,

$$a_{\text{selected}} = \arg\max(\pi(s, a), a) \tag{3.2}$$

here π is the policy. However to ensure that some exploration is made, the ϵ-greedy method is used instead, which means that with probability ϵ a random action is chosen and otherwise the greedy action is used.

The actor's policy would be represented by the dopamine-dependent corticol-striatal projections (Joel et al., 2002). The value function could be found in the orbitofrontal cortex and limbic cortex.

The actor critic architecture is capable of bootstrapping the behaviour without any previous knowledge of the world through trial and error. The original framework of our actor critic model has already been coded (Djurfeldt, 2002).

Our hypothesis states that the action is learned in the basal ganglia then later consolidated in the cortex. In our model we have chosen to implement a labyrinth in a grid world that the basal ganglia model needs to learn the optimal route for (see Section 4). The basal ganglia in our model have four directional commands to move the rat. In addition to these four commands we have implemented a special fifth action. The purpose of this action is to yield control to the cortex model for a step. It is important to note that this is probably not accurate, but it is a first steppingstone for further development of the model.

The motor commands are all issued through the motor cortex model. The difference between the four directional commands and the fifth command in our model can be seen by looking at how the commands are passed to the cortex model. The command is given as a control signal, telling the cortex model where it should head towards, however the fifth action is the absence of such a control signal, which yields control to the cortex model for one step. By associating a small reward with this fifth action the network will be motivated to use it when possible. The idea is that the cortex model learns the new sequence by observing the actions selected by the basal ganglia model. Since it will be more favourable to yield control to the cortex model, provided it gives the correct action, control will be slowly moved from the basal ganglia model to the cortex model, as the network learns the correct sequence of actions to reach the end.

In this model the basal ganglia model makes the choice to yield control to the cortex model or not in each step. In reality this is moved from the basal ganglia to the cortex, freeing up the basal ganglia to learn new things.

3.3   Non-Monotonic Single-Network Morita Model

In this section we study a possible candidate for the model of the cortex. We require that the model should be able to recall a sequence of patterns on receiving a cue. One example of a network capable of recalling sequences is the partial recurrent network, where most of the connections are feed-forward but a selection of feedback neurons is included that allows the network to remember cues from the recent past. The feedback is received through context units. With fixed input a network can be taught to generate a sequence from a selection of sequences (Jordan, 1986, 1989).

Another possible candidate are the attractor networks. Normal attractor networks approach an attractor basin and then get stuck there. In order for them to continue onwards there needs to be some special mechanism. This can be done with adaptation which works by temporarily exhausting the synapses or the neurons, weakening the basin and allowing the network to proceed to the next attractor. A more detailed discussion about adaptation can be found in Sandberg (2003).

The Morita model (Morita, 1996a) is a neural network capable of recalling sequences. Given a cue it can associate between either a pair of patterns or a sequence of patterns, depending on how we choose to train it. The network learns this behaviour by following a learning signal at a distance, and in each step the weight matrix – which holds all the connection strengths between the neurons – is updated in such a way that a flow from the current state to the learning signal is created. This results in the ability to recall patterns from intermediate states and the recall can be done without synchronization.

The model here works with 1000-dimensional non-sparse patterns with equal proportions of ones and minus ones. The instantaneous potential u_i of the i:th cell is calculated by

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^{n} w_{ij} y_j + z_i \tag{3.3}$$

$$y_i = f(u_i) \tag{3.4}$$

where y_i is the output, z_i the external input which is a rescaled version of the learning signal r_i (see below), τ a time constant and n is the number of neurons. The non-monotonic function (Morita, 1996b) is,

$$f(u_i) = \frac{1 - e^{-c_1 u_i}}{1 + e^{-c_1 u_i}} \cdot \frac{1 + \kappa e^{c_2(|u_i| - h)}}{1 + e^{c_2(|u_i| - h)}} \tag{3.5}$$

where we have used the constants $c_1 = 50$, $c_2 = 10$, $h = 0.5$, $\kappa = -1$, see figure 3.1. The learning is performed using a Hebb-like learning rule, the covariance rule, where the recurrent input $Y = (y_1, y_2, ..., y_n)$ is matched with the learning signal $R = (r_1, r_2, ..., r_n)$.

$$\tau' \frac{dw_{ij}}{dt} = -w_{ij} + \alpha r_i y_j \tag{3.6}$$

Note that $\alpha = \phi(y_i)$ depends on y_i. Please note index, if y_j is used instead, then the network recalls the sequence backwards!

$$\phi(y_i) = \begin{cases} 50(0.5 - y_i) & y_i \leq 0.5 \\ 0 & y_i > 0.5 \end{cases} \tag{3.7}$$

The external input vector $Z = (z_1, z_2, ..., z_n)$, which is the pattern shown to the network at a given point in time, is taken to be $Z = \xi R$ where ξ decreases with time.
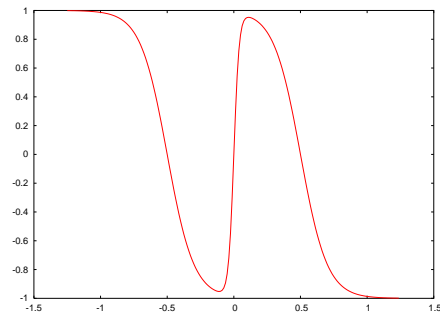
9

Figure 3.1. The non-monotonic function, note that κ decides which values the positive and negative asymptotes converge towards, in our case -1 and +1 respectively.
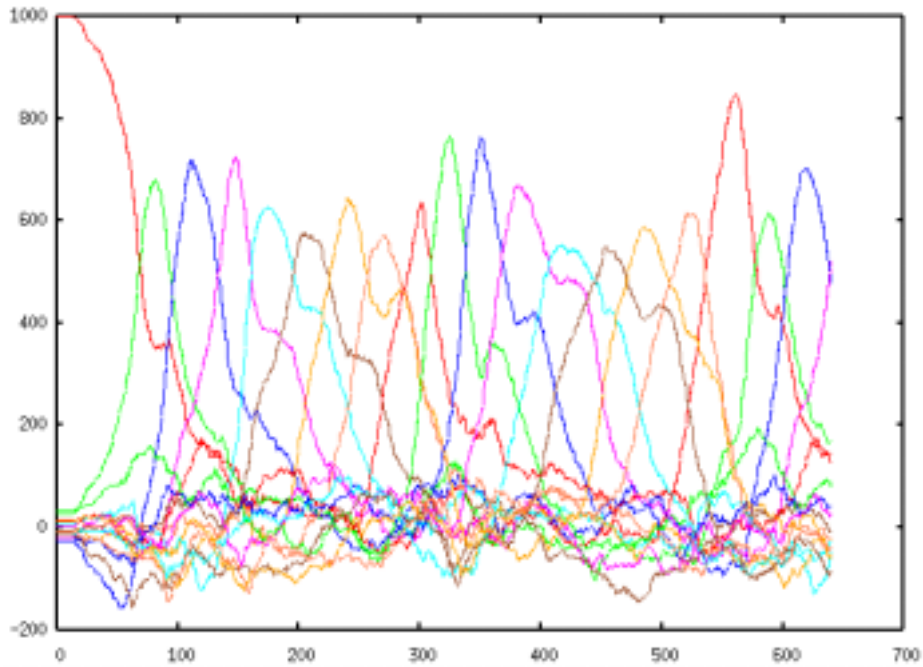

Figure 3.2. The Morita network recalls a sequence of patterns similar to those that will be encountered in our model. Number of steps on the x-axis and overlap on the y-axis. If you view the pdf rather than the postscript version of this report then the above graph might be fuzzy because of the way bitmap images are stored.

10

3.3.1 Results

Using this Morita model we were able to create an association chain of up to sixteen patterns with overlaps of about 60–80%, see figure 3.2. This should be sufficient for our purpose, since we deal with motor sequences twelve patterns long. There was also a speedup of the recall relative to the speed of learning.

3.3.2 Discussion

The most interesting property in the Morita network is the non-monoticity. It ensures a smooth recall of the patterns. Looking at the pattern space, we have a curved surface between consecutive patterns A and B, such that if the network state is far from this surface it travels rapidly towards this almost perpendicular, but as it gets closer the neurons saturate because of the non-monoticity and the network moves more slowly along the curved surface towards pattern B (Morita, 1996b).

The Morita network does not require any special mechanisms during recall, the flow is a natural part of the network in its original configuration. However during learning this version required a set of interpolated patterns between the patterns A and B. These patterns had to be supplied to the network. In section 4.2 we introduce a novel idea for how to remove this dependency on external pattern interpolation.

One concern we had with our implementation of the Morita network was the speedup during recall. Figure 3.3 illustrates this. During learning in the Morita model a flow is created from the current network state towards the learning signal. This flow however is stronger than the flow driving the network forward during learning, which could account for the speedup. This becomes even more obvious when using our alternative learning method as then the flow learned will add to the effect, further increasing the speedup.

There are a couple of other beneficial effects from the non-monotonic transfer functions in addition to the smooth recall, that deserve to be mentioned. Monotonic networks have the problem that the weights keep growing because the weight modifications from the Hebb-like learning rule accumulate over time, this is not the case for non-monotonic networks.

3.4 Dual Network Cell-Pair Morita Model

The biggest difference between the Cell-Pair model and the model described in section 3.3 is that the non-monotonic units have been replaced by pairs of units that together generate the non-monoticity by using monotonic transfer functions (Morita and Suemitsu, 2002; Morita, 1996b). This pair represents a cortical mini-column. The excitatory unit represents a population of pyramidal neurons. The inhibitory unit represents a population of inhibitory interneurons. Sparse pattern coding with 10% ones and 90% zeros have been used here as opposed to non-sparse coding previously used. One of the benefits with sparse coding is the ability to store more patterns.
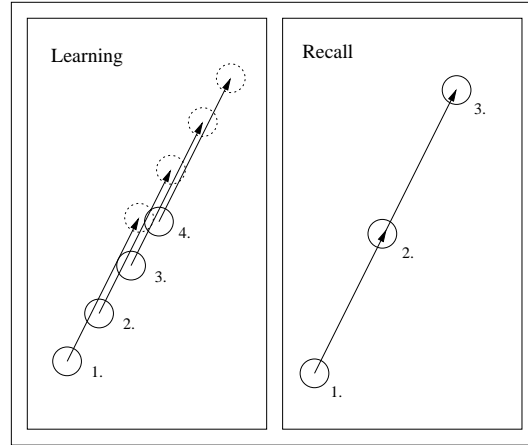
Figure 3.3. The learning mechanism of the Morita network creates a speedup at recall. This effect can be illustrated by looking at how the differential equations are evaluated numerically. During learning, associations are created several steps forward, because of the size of the lag. When recalling some of the intermediate steps originally learned will be skipped over.

The Dual Network model consists of two parts: an interpolating network and a learning network. The first one receives the input, and is responsible for creating a slow moving learning vector that moves from the pattern A to the pattern B. It does so by having a competing network, when the first pattern is shown the network relatively quickly tunes to it and when the second pattern is displayed the competition will create a slow movement from the old to the new pattern. The second network is similar in function to the model described in section 3.3. It is responsible for learning and recalling the sequences.

Unfortunately there were some issues with getting this version of the Morita model to work properly. The network failed to recall the learned patterns, sometimes remaining completely static. In retrospect this might have been because of the random connection weights used to and from the network not being properly chosen.

3.4.1 Interpolating Network

The purpose of the interpolating network is to create a learning signal for the other network to learn from. When the simulation starts, there is no activity in the interpolating network, and the first pattern displayed will quickly determine the activity in the network. Then follows the delay phase, after which a target pattern is displayed, but because there is some activity already present in the interpolating network the change is not instantaneous. Instead, what we get is a natural interpolation between the representations of the cue and target patterns. The network also receives input from the learning network.

In order to create an interpolation between the two patterns the network also has to apply random weight matrixes to the input signal and the recurrent feedback.

Table 3.1. General Network Parameters for the Dual Network Morita Model

θ	$=$	3	β_1	$=$	25	ρ	$=$	0.016
w_i^*	$=$	10	β_2	$=$	50	σ	$=$	0.8
λ	$=$	0.3	γ	$=$	0.05	c	$=$	10
τ'	$=$	50000τ				η	$=$	$\begin{cases} 0 & \text{normally} \\ 0.75 & \text{delay phase} \end{cases}$

The situation is similar to the one discussed in section 4.2. If several patterns in an intermediate sequence of patterns only differ in amplitude then recall will fail. The equations for updating the interpolating network are,

$$\tau\frac{d\nu_i}{dt} = -\nu_i + \sum_{j=1}^{n} p_{ij}s_j + \sum_{j=1}^{n} q_{ij}x_j - \rho\sum_{j\neq i} r_i + \sigma r_i + \eta \qquad (3.8)$$

$$r_i = f(v_i) \qquad (3.9)$$

The function $f(u)$ is a monotonic sigmoid function given by

$$f(u) = \frac{1}{1 + e^{-cu}} \qquad (3.10)$$

The values used for p and q were taken from two normal distributions, one with an average value of $5.5 \cdot 10^{-3}$ and variance $3.7 \cdot 10^{-4}$ and the other one with mean value $7.2 \cdot 10^{-3}$ and variance $6.4 \cdot 10^{-4}$. Both normal distributions were cut off at three standard deviations to avoid extreme values.

3.4.2 Learning Network

The learning itself is performed in a separate network, here referred to as the learning network. It uses a Hebbian learning rule, with time constants considerably lower than the other time constants in the model. The first equation below handles the weight matrix for the excitatory neurons in each pair. The second equation handles the weight matrix for inhibitory neurons.

$$\tau'\frac{dw_{ij}^+}{dt} = -w_{ij}^+ + \alpha r_i x_j \qquad (3.11)$$

$$\tau'\frac{dw_{ij}^-}{dt} = -w_{ij}^- - \beta_1 r_i x_j + \beta_2 x_i x_j + \gamma \qquad (3.12)$$

Here α, β_1 and β_2 are learning coefficients, and γ represents the lateral inhibition among units.

The excitatory and inhibitory units of the network are updated according to,

$$y_i = f(\sum_{j=1}^{n} w_{ij}^- x_j - \theta) \tag{3.13}$$

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^{n} w_{ij}^+ x_j - w_i^* y_i + z_i \tag{3.14}$$

$$x_i = f(u_i) \tag{3.15}$$

where x_i and y_i are the outputs of the excitatory and inhibitory units respectively, u_i is the potential, z_i the external input, w_{ij}^+ and w_{ij}^- the synaptic weights for the excitatory and inhibitory units.

3.4.3 Discussion

The model described in section 3.4 uses sparse population coding with 10% ones and the rest zeros. The sparse coding increases the capacity of the network, so that more patterns can be stored given a fixed number of neurons. In the computer simulations we performed, a network size of 1000 pairs of neurons was used. The parameters used can be found in table 3.1.

Unfortunately time constraints prevented the full implementation of this version of the model. Instead we chose to investigate a spiking model (section 3.5) to see what would be required for the next phase, since our goal was to be able to compare the simulated data with data recorded from biological models, and since data from a spiking neuron model would be better suited for comparison with recorded spiking data from rats.

3.5 Spiking Model

As mentioned above we are trying to replicate the experimental conditions in Jog et al. (1999). This will serve as a foundation for comparison between our simulated neurons and their real biological counterpart. The data from the experiments are spiking, but neither of the two models discussed in previous sections have spiking neurons, instead they deal with rates. Eventually we hope to move from a rate based cortex model to a spiking model capable of both learning and retrieving sequences in the presence of distractors. A spiking neural model would allow for better comparison. The network also needs to be able to retain learned memories despite applied distractors during delay phase before recall.

In this section we study a spiking neural network model with pyramidal cells and interneurons capable of recalling patterns, and retaining them in memory despite being exposed to distractor patterns of strength equal to the original input (Brunel and Wang, 2001). The key feature in this model is the domination of inhibition.

This model is different from the previous cortex models already discussed as it has no mechanism for learning new patterns, however the fact that it uses spiking

neurons as well as the ability to retain memory despite external distractors makes it interesting to study.

Another important difference between the rate models and the spiking model discussed here is that the former store the memories in the synapses and the latter uses the activity in the cell assemblies. It is important to note that the spiking model implemented here does not have any learning, instead the weight matrix is fixed. It receives binary patterns through external neurons that fire randomly with small variations between the different elements. The pattern is encoded as a variation of probability of firing, where a one is encoded as a slightly higher firing rate for that element than for an element coding for a zero.

Short term memories are stored by activity in cell assemblies of pyramidal cells which code for a particular pattern. They retain the activity by having stronger connections to cells within the cell assembly than to those outside it. At the same time their spikes trigger the firing of the interneurons that place a general inhibition on all pyramidal cells. This inhibition suppresses activity in other cell assemblies and makes the network resistant to the distractor patterns. Since there is a pattern already present, any new patterns have to fight the inhibition from the interneurons to establish themselves. On the other hand, the first pattern displayed can quickly establish itself because the inhibiting interneurons are not activated by any previous patterns.

The strong inhibition also gives a neat way of clearing the network from old patterns. By activating all external neurons the network receives a strong input which leads to a strong inhibitory response from the interneurons that clears out all activity in the pyramidal cells.

3.5.1 External Neurons

The external neurons represent the input to the network. They randomly fire at about 3 Hz, which corresponds to real neurons in the cerebral cortex. By variations in the frequency they fire, binary patterns can be coded, where a slightly higher rate of firing represents a one, and a lower rate codes for a zero. Each one of the external neurons contributes with on average three spikes per second, and together the 800 external neurons send data to the network at a rate of 2.4 kHz.

3.5.2 Pyramidal Cells

Most cells that make up this model and where the memory activity is retained are the pyramidal cells. These cells form dense excitatory connections to one another, where the strength of weights are dependent on whether the cells belong to the same or a different cell assembly. Focused persistent activity can be retained in cell assemblies by having stronger connections between cells within the assembly than cells outside, together with a general inhibition from the internal neurons due to the firing of the active cell assembly. Every cell assembly contains $r_{cs}N_E$ neurons, where N_E is the total number of excitatory pyramidal cells, and each of the p cell

Table 3.2. Pyramidal Cell Constants, please note that the membrane time constant τ_m is C_m/g_m

Resting Potential	V_L	$=$	-70 mV
Firing Threshold	V_{thr}	$=$	-50 mV
Reset Potential	V_{reset}	$=$	-55 mV
Membrane Capacitance	C_m	$=$	0.5 nF
Membrane Leaky Conductance	g_m	$=$	25 nS
Refractory Time	τ_{ref}	$=$	2 ms
Membrane Time Constant	τ_m	$=$	20 ms
Number of Patterns	p	$=$	5
Relative Cell Assembly Size	r_{es}	$=$	0.1

assemblies codes for different sparse patterns. The remainder of the pyramidal cells code for no specific pattern.

The individual pyramidal cells are modelled as leaky integrate-and-fire neurons (Tuckwell, 1998). Below threshold the membrane potential $V(t)$ is calculated by

$$C_m \frac{dV(t)}{dt} = -gm(V(t) - V_L) - I_{\text{syn}}(t) \tag{3.16}$$

where $I_{\text{syn}}(t)$ is the total synaptic current flowing into the cell. The pyramidal cell constants can be found in table 3.2.

3.5.3 Interneurons

The spiking model is dominated by inhibition. This inhibition is created by 200 interneurons modelled as leaky-integrate-and-fire neurons in analogy with the pyramidal cells, please note that some of the constants differ, see table 3.3.

3.5.4 Synapses

AMPA and NMDA are receptors on excitatory synapses, and GABA inhibitory. The external neurons use AMPA. Pyramidal cells use both AMPA and NMDA, the latter one has a slower response time and longer time scale and plays a vital role in retaining persistent activity within the pyramidal cell assemblies. GABA is used exclusively by the interneurons in this model. All synapses have a latency of 0.5 ms.

The total synaptic currents are given by

$$I_{\text{syn}}(t) = I_{\text{AMPA,ext}}(t) + I_{\text{AMPA-rec}}(t) + I_{\text{NMDA,rec}}(t) + I_{\text{GABA,rec}}(t) \tag{3.17}$$

Table 3.3. Constants for Interneurons, values that differ from pyramidal cells are in italic

Resting Potential	V_L	$=$	-70 mV
Firing Threshold	V_{thr}	$=$	-50 mV
Reset Potential	V_{reset}	$=$	-55 mV
Membrane Capacitance	C_m	$=$	0.2 nF
Membrane Leaky Conductance	g_m	$=$	20 nS
Refractory Time	τ_{ref}	$=$	2 ms
Membrane Time Constant	τ_m	$=$	10 ms

where

$$I_{\text{AMPA,ext}}(t) \quad = \quad g_{\text{AMPA,ext}}(V(t) - V_E) \sum_{j=1}^{C_{\text{ext}}} s_j^{\text{AMPA,ext}}(t) \tag{3.18}$$

$$I_{\text{AMPA,rec}}(t) \quad = \quad g_{\text{AMPA,rec}}(V(t) - V_E) \sum_{j=1}^{C_E} w_j s_j^{\text{AMPA,rec}}(t) \tag{3.19}$$

$$I_{\text{NMDA,rec}}(t) \quad = \quad \frac{g_{\text{NMDA}}(V(t) - V_E)}{1 + [Mg^{2+}]e^{-0.062V(t)}/3.57} \sum_{j=1}^{C_E} w_j s_j^{\text{NMDA}}(t) \tag{3.20}$$

$$I_{\text{GABA,rec}}(t) \quad = \quad g_{\text{GABA}}(V(t) - V_I) \sum_{j=1}^{C_I} s_j^{\text{GABA}}(t) \tag{3.21}$$

here $V_E = 0$ mV is the reversion potential for excitatory cells and $V_I = -70$ mV for inhibitory cells. Charged ions block the membrane, reducing the ion flow, to account for this we use the Jahr and Stevens' (Jahr and Stevens, 1990) formula for modulation of NMDA currents by the extra cellular magnesium concentration where $[Mg^{2+}] = 1$ mM. The weights w_j that appear above will be described further down.

The network is recurrent, which means that all cells have connections to one another. When we talk about external synapses, we mean those originating from cells outside the network. To measure the synaptic activity we use s, which represents the fraction of open channels. For both external and recurrent AMPA, we have

$$\frac{ds_j^{\text{AMPA}}(t)}{dt} = -\frac{s_j^{\text{AMPA}}(t)}{\tau_{\text{AMPA}}} + \sum_k \delta(t - t_j^k) \tag{3.22}$$

where $\tau_{AMPA} = 2$ ms. The sum over k is a sum over all the spikes emitted by the presynaptic neuron j. The input spikes from the external neurons are modelled by Poisson processes with the rate ν_{ext}, and are independent for each post synaptic cell.

17

Table 3.4. Recurrent Synaptic Conductances in nS

	Pyramidal cells	Interneurons
$g_{\text{AMPA,ext}}$	2.08	1.62
$g_{\text{AMPA,rec}}$	0.104	0.081
g_{NMDA}	0.327	0.258
g_{GABA}	1.25	0.937

In other words, two cells do not normally receive the same input from the external neurons. For the slower NMDA channels we use

$$\frac{ds_j^{\text{NMDA}}(t)}{dt} = -\frac{s_j^{\text{NMDA}}(t)}{\tau_{\text{NMDA,decay}}} + \alpha x_j(t)(1 - s_j^{\text{NMDA}}(t)) \tag{3.23}$$

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_{\text{NMDA,rise}}} + \sum_k \delta(1 - t_j^k) \tag{3.24}$$

where $\tau_{\text{NMDA,decay}} = 100$ ms, $\alpha = 0.5 ms^{-1}$ and $\tau_{NMDA,rise} = 2$ ms. The NMDA receptors are critical for the retention of patterns in the pyramidal cells, because of the slower time constants and longer decay time for NMDA that helps keep the activity high within the population. For the inhibitory GABA we have

$$\frac{ds_j^{\text{GABA}}(t)}{dt} = -\frac{s_j^{\text{GABA}}(t)}{\tau_{\text{GABA}}} + \sum_k \delta(t - t_j^k) \tag{3.25}$$

where $\tau_{\text{GABA}} = 10$ ms. The values for the recurrent synaptic conductances (see table 3.4) are about 1 nS in magnitude and correspond roughly with the experimental data.

3.5.5 Connectivity

The pyramidal cell population is divided into cell assemblies that each codes for a different pattern. In this version of the model the weights are hard coded, but represent what Hebbian learning would achieve; neurons that correlate often have a stronger weight than neurons that do not correlate. Here this can be seen by connections between neurons within the same cell assembly having a stronger weight than connections that are between neurons in different cell assemblies, see figure 3.5.5.

We set $w_j = w_+$ when both neurons code for the same pattern. Between neurons coding for different patterns, and from the non-selective neurons to a selective neuron, we set $w_j = w_-$. All other weights w_j are set to 1. With $w_+ = 2.1$ and w_- calculated from $w_- = 1 - f(w_+ - 1)/(1 - f)$ the overall synaptic drive is kept constant as w_+ varies (Amit and Brunel, 1997). The network is fully connected, meaning all neurons connect to all other neurons.
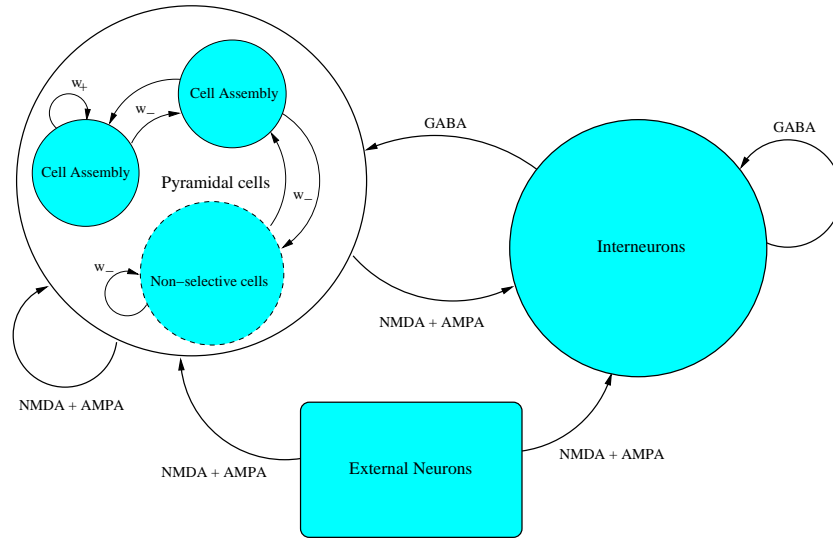
18

Figure 3.4. Neuron connections in the Brunel model.

3.5.6 Results and Discussion

The simulation was run with a time step of 0.01 ms using a second order Runge-Kutta method. To increase accuracy the spike times were interpolated within each time step (Hansel et al., 1998).

Figure 3.5 shows the problem with having too strong input signals. Cue and distractor patterns all have the same strength. If it is too high, the distractor will drown the activity in the network. If, on the other hand, the strength is too low (figure 3.6) then the cue will not overcome the initial inhibition barrier and fail to establish any persistent activity in the network.

Figure 3.5. Brunel simulation with too strong input relative to the inhibition from the interneurons. The distractor erases the original pattern from memory. The x-axis shows time, and the y-axis activity. The input strength was multiplied by λ which here was 0.0375. At step 100 the pattern is shown for the network, then turned off at 200. At 250 the distractor is activated and soon dominates the network.
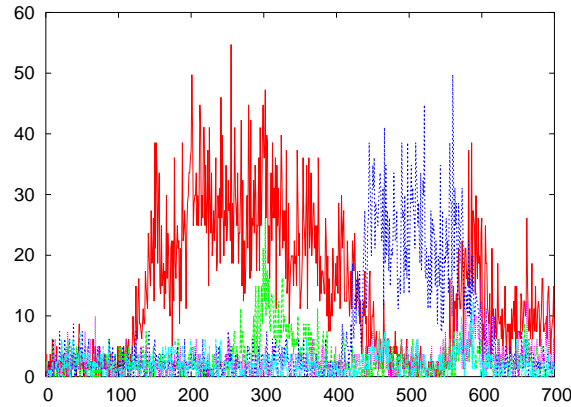

Figure 3.6. Spiking simulation with $\lambda = 0.0295$. Here the original pattern survives the first distractor, but the second one clears it from memory. Note that the strength of the original pattern and the distractors is the same. Here again the pattern is shown at step 100, the activity takes longer to establish itself in the cell assembly than for higher λ values, see figure 3.5. The first distractor is unable to overcome the inhibition but the second distractor manages to establish itself.

Chapter 4

Combined Cortex-Basal Ganglia Model

Experiments by Matsumoto et al. (1997) on a behaving monkey taught to press a sequence of buttons indicate that a unilateral lesion of the Striatum/Substantia Nigra Compacta (SNc) prior to learning affects how the learned motor sequence is stored. When the monkey was given the reward prior to pressing the last key it altered its behaviour. However a reference monkey that received the lesion after learning was unable to adapt to the change and kept pressing the last button by habit. This difference in behaviour suggests that in the monkey receiving the lesion prior to training the sequence was not chunked together into a macro, since it was able to learn to use the shorter sequence despite the lesion. However in the reference monkey the behaviour sequence was stored as a unit and after the lesion it was incapable of changing it.

In this chapter our aim is to describe the model that we use to simulate the interaction between the basal ganglia and the cortex during learning of sequential behaviour in the brain (see hypothesis in previous chapter). Here the basal ganglia are modelled with a network based on Sutton and Barto (1998) actor-critic architecture, and the cortex is modelled by the recurrent attractor network described in section 3.3.

This model in turn controls a simulated rat located in a grid world, see figure 4.1. Here, we are trying to replicate the experimental conditions in Jog et al. (1999), where a rat is placed in a labyrinth shaped like a 'T'. There are two exits, one on each side of the intersection. A tone is played for the rat when it is half way through the labyrinth indicating if it should turn left or right. If the rat reaches the correct end it is rewarded, otherwise no reward is given.

The model can be broken down into two natural parts: the basal ganglia model and the cortex model. Sensations from the world are received by the sensory cortex model, see figure 4.2. Each location in the grid world is associated with a unique sensation. The rat also receives an audio tone. Both are coded as random patterns. For more details on how this is done see Appendix A. From the sensory cortex model the sensations are passed on both to the basal ganglia model and the motor cortex model. The basal ganglia model evaluates the situation and sends its command to the motor cortex model. The motor cortex model issues motor commands, and also
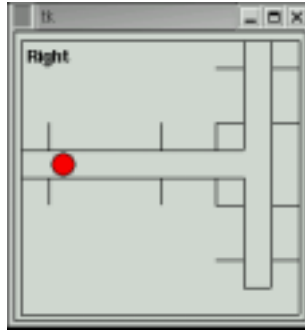
Figure 4.1. The rat, represented as a dot, during one of the simulations.

feeds back to the basal ganglia model.

The thalamus has been omitted in this model. We assume it is responsible for controlling attention span. The other parts can find their representation in the actor critic method used for learning, see section 4.1.

4.1 Basal Ganglia

The basal ganglia model was based on the actor-critic architecture from Sutton and Barto (1998), a brief introduction can be found in section 3.2.

The basal ganglia model are involved in learning the best way through the labyrinth. It receives input from the sensory cortex model and at its disposal it has four movement commands, one for each direction. This is enough to both learn and control the rat. However what we are interested in is a mechanism that can, in a natural way, transfer control from the basal ganglia model to the cortex model. To do this we propose the fifth action, as discussed earlier in section 4.2, which for one step yields control to the cortex model. We associate a small reward with this fifth action, so that it will be slightly more rewarding to use the cortex model instead of the basal ganglia model. The idea is that once the cortex model has learned to give the correct motor commands, the basal ganglia model will yield more and more to the cortex model. In other words, control is transferred from the basal ganglia model to the cortex model. One way of putting this is that the basal ganglia model is only responsible for correcting the behaviour of the cortex model. Before the cortex model has learned the correct sequence of actions to navigate through the labyrinth there are a lot of actions to correct, but as the cortex model gets progressively better the basal ganglia model performs less corrections and becomes more and more quiet.

4.2 Cortex

The cortex model is based on the Morita network discussed in section 3.3. The task for this network is to learn the correct action sequences by learning from the basal
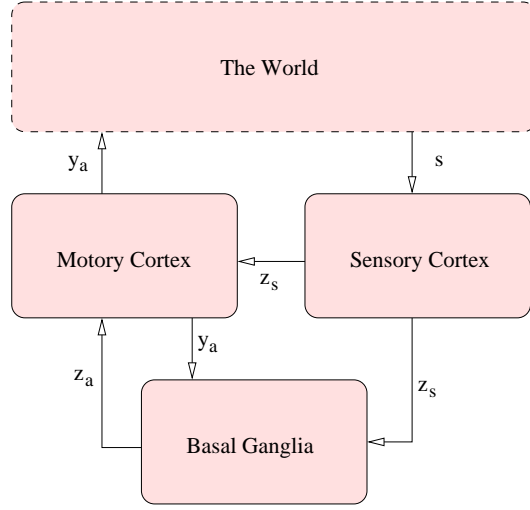
Figure 4.2. Information flow in our prototype. 1. The sensory cortex model receives latest sensation from the world. 2. The basal ganglia model evaluates the sensation and the last action and decides on a new action. 3. Motor cortex model receives a new action, this action command can also be the special "no signal, let cortex decide alone". 4. The motor cortex model creates an action that is sent back to the world.

ganglia model. There are alternatives to this model, which are also described in that section.

Two versions were implemented, one where all of the basal ganglia model's actions were issued through the cortex model, and a second model where the basal ganglia model issued the actions directly. In both cases the cortex received the action chosen by the basal ganglia model, but in the second case it could not alter it. The latter can be motivate by Parent et al. (2001) which indicates that basal ganglia connect directly to the lower parts of the motor system. A copy of the command is set through collaterals to the thalamus.

As mentioned in section 4.1 the basal ganglia model has a fifth action that temporarily yields control to the cortex model. Normally the basal ganglia model tells the cortex model the new action, and this is used as a learning signal which the network slowly shifts towards. This creates an association between the previous state and the new one. When the fifth action is issued the cortex model only receives the sensory input and has to decide the action on its own.

The Morita model uses interpolated patterns to create this smooth flow. This is not very biological. To replace this and eliminate the interpolation we propose a novel idea which is capable of creating a similar flow between cue and target patterns.

We propose the usage of an alpha function that is a delayed version of the network state and which follows the current network state at a distance, see figure 4.3. The weight updating rule uses this alpha function and the network's current state as the two reference points when creating the association and updating the weights.
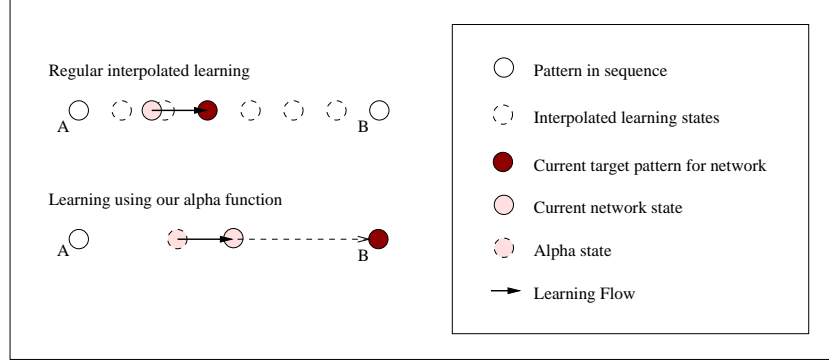
23

Figure 4.3. Interpolation and our proposed alpha function learning. The alpha state is a time lagged version of the network state, learning is performed from the old state to the new, creating a flow towards the goal.

The alpha function, here denoted by a, is updated as follows:

$$\tau_a \frac{db}{dt} = -b + f(u) \tag{4.1}$$

$$\tau_a \frac{da}{dt} = -a + b \tag{4.2}$$

where u is the instantaneous potential of the network, τ_a a time constant that decides the response time of the alpha function, i.e. how much it lags behind the current network state. By using the alpha function we are able to avoid interpolations while still using the Hebbian type learning rule to update the weight matrix,

$$\tau' \frac{dw_{ij}}{dt} = -w_{ij} + \alpha \theta(u_i) a_j \tag{4.3}$$

$$\theta(u) = \frac{1 - e^{-50u}}{1 + e^{-50u}} \tag{4.4}$$

Here θ is a threshold. In order to make recall possible we must prevent the network from having two neighbouring intermediate patterns only differ by a scale factor, because if that happened an attractor would be created, resulting in the network becoming static during recall. Our solution is to randomize the time constants slightly. This ensures that the switching of the elements will be spread out in time. The new updating rule looks like this:

$$\tau_i \frac{du_i}{dt} = -u_i + \sum_{j=1}^{n} w_{ij} y_j + z_i \tag{4.5}$$

$$y_i = f(u_i) \tag{4.6}$$

where τ has been replaced by τ_i, which varies between neurons.

Chapter 5

Results

Figure 5.1 comes from an early version of the model before the reward for cortex action was implemented. It shows the overlap between the optimal action sequence for a rat running through the labyrinth and the network states used after completed learning. Each peak in the graph represents one step. This graph shows that the basal ganglia model is able to learn the optimal route.

Figure 5.2 is a trial run, where the cortex malfunctioned after 55 iterations. The reason for this is probably the speedup during recall relative to the speed when learning, which causes timing issues in combination with cortex weights that had become so strong that they overruled the action commands given by the basal ganglia.

We mentioned in section 4.2 that there were two different versions of the cortex model implemented, one where all the motor commands from the basal ganglia model went through the cortex model before taking action, and a second where the basal ganglia model directly issued motor commands, except when using the fifth action and yielding in favour of the cortex model.

The third figure 5.3 shows a test run where the cortex model has been prevented from overruling the basal ganglia model. In other words, the cortex model's output is only used if the basal ganglia model yields control to the cortex model, as opposed to having all the output from the basal ganglia model go through the cortex model as in the original situation. Note the second bump in the graph after which there are no or very few requests for cortex actions. The cortex is no longer reliable, and the basal ganglia model is issuing almost all motor commands. Four of the runs in the figure reached the wrong goal, this happened during trial runs 2, 14, 66 and 70. The last two followed shortly after the second bump and emphasized the fact that the network had to reevaluate the value of the cortex actions.

Figure 5.4 shows the number of times the cortex model was called divided by the number of times the basal ganglia model made the final decision plotted against the number of trials. The initial bump in the graph is because it takes some time for the rewards from reaching the goals to propagate through the value function. During that time having the network listen to the cortex model and getting the small reward associated with it is favourable. Prior to the second bump the basal ganglia model has learned that listening to the cortex model is again a good strategy, but what
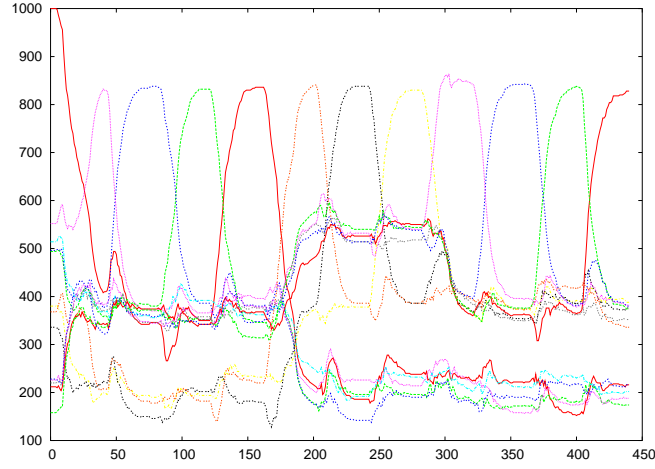
Figure 5.1. Overlap between patterns corresponding to the optimal action and network states, no extra reward was given when an action was selected by the cortex model. Overlap is shown on the y-axis and time steps on the x-axis.
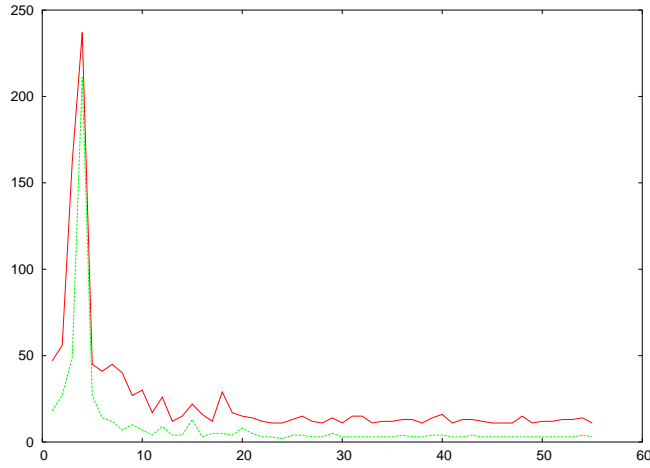

Figure 5.2. This is a plot of the number of steps taken before one of the two goals was reached overlaid with the number of times the basal ganglia model was silent letting the cortex model decide the action alone and both plotted against the 55 first trials. As you can see the percentage of cortex actions are relatively high initially, this is because the goal rewards have not yet propagated through the value function, and using the cortex model's command gives a small reward which at that point in time looks favourable, and thus gets more usage than merited.
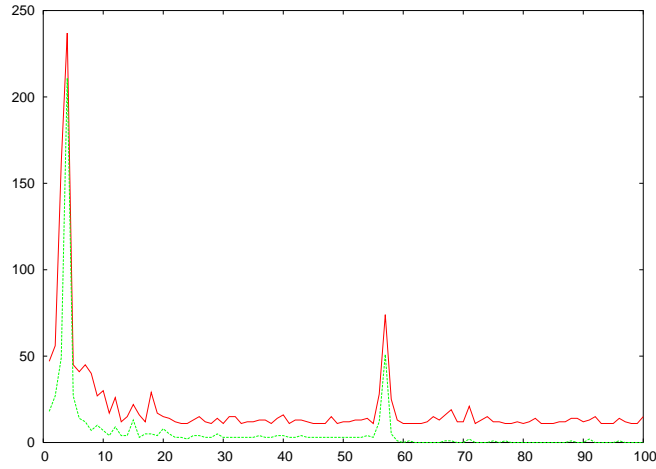
Figure 5.3. This figure shows the number of steps used to reach a goal and number of times the cortex model was used plotted versus number of trials. Time steps are on the x-axis. The second bump in the graph is when the matrix becomes unreliable. This might be a timing issue, since playback at recall is faster than the learning speed.

happens next is probably due to the speedup of recall in the cortex model, as the weights appear to be strong enough to partly ignore the control signal from the basal ganglia model. This situation worsens as the activity in the network is built up over several iterations. The rat tries to run up or down towards the goal, as it should do at the intersection, but since it is not yet there it gets stuck. After a few trials with intense activity the cortex model usage goes down, its output is no longer reliable. This is very unfortunate, but it should be possible to reduce or perhaps eliminate most of the lag induced speedup by careful adjustment of the time constants.
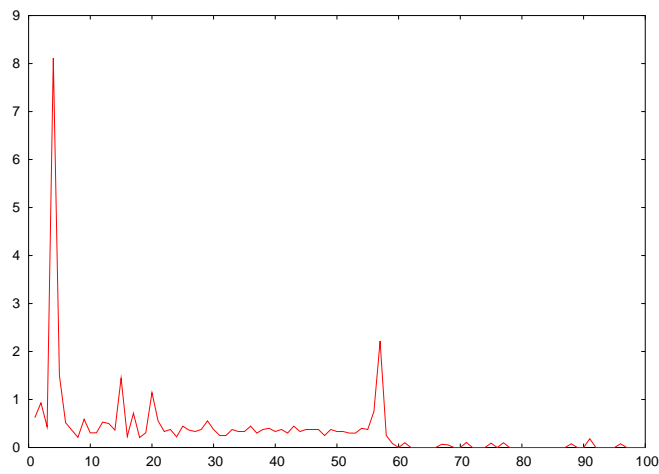
Figure 5.4. Usage of the cortex model plotted relative usage of the basal ganglia model. Note the two bumps which indicate changes in the behaviour of the network. The first bump is due to initial exploration, before the goal reward has propagated through the value function. The second bump is due to the recall speedup of the cortex model, which makes it unreliable. Note the low activity in the cortex model after the bump.

Chapter 6

Discussion

The basal ganglia are involved in the learning of motor behaviour; the common view is that action sequences are learned in cortex then moved down into the basal ganglia as they become habitual. Using the reinforcement learning network itself to store learned behaviour, and have update mechanisms silent, seems like a waste of resources. Instead we believe that the behaviour is learned by the basal ganglia but later consolidated in the cortex. An interesting question that we ask is, how could this be done?

In this report we try to give a plausible model with a network that is capable of both learning a behaviour and a mechanism for transferring it to a secondary network for storage and retrieval.

There are other structures of the brain that also have a similar architecture. One hypothesis is that the Medial Temporal Lobe (MTL) stores a snapshot through fast learning and then later the memory is consolidated in neocortex. After repeated reactivation the neocortex becomes capable of performing the memory retrieval without the MTL, (Gluck et al., 1997; Granger et al., 1996).

In figure 5.1 it is interesting to note the increased noise level in the mid part of the graph, this is the part of the labyrinth where the sound indicating direction has been turned on, but before the rat has been allowed to turn. This could be interpreted as a conflict between behaviours on what to do, the rat is supposed to run forward while at the same time the audio signal tells it to turn left or right. Further information about the audio, location and action parts of the patterns can be found in Appendix A.

The problems with the cortex model during later parts (figure 5.2) of the learning is probably due to a speedup relative to the speed at which the sequence was originally presented by the basal ganglia model. To have the same speed during both recall and learning it is important that $\frac{dr}{du} = 1$ but our alpha function required $\tau_\alpha = 1.1$ which results in the speedup. The alpha function has two time constants that both are labelled τ_α, perhaps by distinguishing between the two this effect could be reduced. Another alternative is to have some sort of synchronization mechanism to prevent the cortex model from getting too far ahead.

The overlap for a pattern starts to increase while the previous pattern is still

active. This feature has been seen in (Averbeck et al., 2002) where the relative order of actions in a sequence could be found by studying the relative activation in neurons corresponding to those actions.

There are other models for the cortex-basal ganglia system. Nakahara et al. (2001) used a different architecture to model the "2x5 task" on monkeys. In that experiment a monkey is shown a five by five grid where two squares lit up. The task for the monkey is to press them in the correct order, which it has to learn by trial and error. Several of these grid configurations are shown one after another. Some sequences remain the same over days, while others change daily, allowing for comparison between long and short term learning. Their model has two parallel systems, one fast learning visual loop with temporary working memory and quick reset times and a second slower loop, called the motor loop, which takes longer to learn but can recall learned sequences faster. These two loops are then coordinated by the presupplementary motor area (pre-SMA). In the beginning of learning the coordinator uses the visual loop mainly, but as learning progresses the motor loop gets more responsibility.

With this model they have found interesting correlations between simulated training data and data measured from real monkeys during training. Reversed sequences were as hard to learn as new ones and using the opposite hand for the trial on a learned sequence increased the errors, but there were still fewer errors than on a completely new sequence. They also tried blocking parts of the modules in the network and the results correlated with what they expected from a real world experiment.

However the experiments performed by Matsumoto et al. (1997) on monkeys mentioned in the beginning of this chapter favours our model layout over this.

Chapter 7

Conclusions

In our modified Morita model the implementations of the alpha function and randomized time constants were a success and allowed us to replace the artificial interpolation with a more natural mechanism. There is still some work required to create a successful lasting transition of the learned behaviour from the basal ganglia model to the cortex model, because the interaction between the two is complex, and altering one parameter affects both networks. We are however confident that this can be done.

Future models will hopefully be a hybrid between the ideas of Morita, Brunel and our prototype model. Spiking neurons is something that would be interesting to implement in order to allow us to compare data from simulations with biological data from real rats. Given more time the problems due to speedup in the recall using the Morita model should be solvable.

7.1   Acknowledgements

Appendix A

Pattern Coding

The cortex model uses non-sparse patterns with equal parts of ones and minus ones. The basal ganglia model on the other hand need sparse binary patterns in order to work properly. Since the two need to exchange data the patterns had to be translatable during run time, this is done by the simple transformation

$$f(x) = \frac{x+1}{2} \tag{A.1}$$

which fixes the range, but does not translate a sparse pattern to a non-sparse pattern, since they contain different proportions of ones.

In order to solve the problem with non-sparse to sparse conversion we divided the patterns into two parts A and B. The first part, A, was sparse with 10% ones and the rest minus ones. The second part, B, which was three times the size of A, had just over 63% ones and the rest minus one. When both A and B are taken together the complete pattern has the proportions required by a non-sparse network, with half the elements ones and the rest minus ones. A taken alone functions as a sparse representation of the pattern.
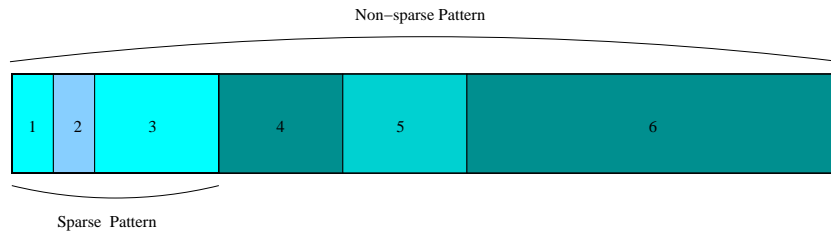

Figure A.1. Encoding of patterns in sparse and non-sparse parts. 1 is audio, 2 action, 3 location, 4 audio, 5 action, 6 location. Here 4, 5 and 6 compensate for the sparseness of 1, 2 and 3 to make the overall pattern non-sparse.

Our original approach was to have A and B of equal size where A had 10% ones and B had 90% of its elements containing ones. This created fix point attractors and a static network. The reason for this is that given two random patterns α and

β, if we look at part B 90% of the elements are ones, which means that many of the ones will be static when going from α to β, and similarly for part A. This builds up fix point attractors and the network is unable to learn associations because 82% of the elements remain the same compared to 68% in our case.

This division of patterns into sparse and non-sparse parts does not have a biological analogy, it is simply a way to merge the two networks together. Later versions of the prototype will have sparse coding throughout and then this special pattern structure will not be necessary.

Bibliography

D. J. Amit and N. Brunel. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex*, 7:237–252, 1997.

B. B. Averbeck, M. V. Chafee, D. A. Crowe, and A. P. Georgopoulos. Parallel processing of serial movements in prefrontal cortex. *Proc Natl Acad Sci U S A.*, 2002.

J. Brown, D. Bullock, and S. Grossberg. How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *Journal of Neuroscience*, 1999.

N. Brunel and X. J. Wang. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, 11:63–85, 2001.

J. L. Contreras-Vidal and W. Schultz. A predictive reinforcement model of dopamine neurons for learning approach behaviour. *Comparative Neuroscience*, 1999.

M. Djurfeldt. Rat-in-Maze framework for See, 2002.

M. Djurfeldt, Ö. Ekeberg, and A. M. Graybiel. Cortex-basal ganglia interaction and attractor states. *Neurocomputing*, (38-40):573–579, 2001.

M. A. Gluck, B. R. Ermita, L. M. Oliver, and C. E. Myers. Extending models of hippocampal function in animal conditioning to human amnesia. *Memory*, 1997.

R. Granger, S. P. Wiebe, M. Taketani, and G. Lynch. Distinct memory circuits composing the hippocampal region. *Hippocampus*, 1996.

A. M. Graybiel. The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, 70:119–136, 1998.

D. Hansel, G. Mato, C. Meunier, and L Neltner. On numerical simulations of integrate-and-fire neural networks. *Neural Comput.*, 1998.

J. C. Houk, J. L. Adams, and A. G. Barto. *Models of Information Processing in the Basal Ganglia*, chapter A model of how the basal ganglia generate and use

neural signals that predict reinforcement, pages 249–270. M.I.T. Press, Cambridge, U.S.A., 1995.

C. E. Jahr and C. F. Stevens. Voltage dependence of nmda-activated macroscopic conductances predicted by single-channel kinetics. *The Journal of Neuroscience*, 10(9):3178–3182, September 1990.

D. Joel, Y. Niv, and E. Ruppin. Actor-critic models of basal ganglia: new anatomical and computational perspectives. *Neural Networks*, 2002.

M. S. Jog, Y. Kubota, C. I. Connolly, V. Hillegaart, and A. M. Graybiel. Building neural representations of habits. *Science*, 286:1745–1749, 1999.

M. I. Jordan. Attractor dynamics and parallelism in a connectionist sequential machine. In *Proceedings of the Eight Annual Conference of the Cognitive Science Society*, 1986.

M. I. Jordan. A parallell distributed processing approach. In *Advances in Connectionist Theory: Speech*. Hillsdale: Erlbaum, 1989.

N. Matsumoto, T. Minamimoto, A. M. Graybiel, and M. Kimura. The effects of unilateral nigrostriatal dopamine depletion on the learning and memory of sequence motor tasks in monkeys. submitted for publication, 1997.

M. Morita. Computational study on the neural mechanism of sequential pattern memory. *Cognitive Brain Research*, 5:137–146, 1996a.

M. Morita. Memory and learning of sequential patterns by nonmonotone neural networks. *Neural Networks*, 9(8):1477–1489, 1996b.

M. Morita and A. Suemitsu. Computational modeling of pair-association in inferior temporal cortex. *Cognitive Brain Research*, 13:169–178, 2002.

H. Nakahara, K. Doya, and O. Hikosaka. Parallel cortico-basal ganglia mechanisms for acquisition and execution of visuomotor sequences – a computational approach. *Journal of Cognitive Neuroscience*, 2001.

A. Parent, M. Lévesque, and M. Parent. A re-evaluation of the current model of the basal ganglia. *Parkinsonism and Related Disorders*, 7:193–198, 2001.

A. Sandberg. *Bayesian Attractor Neural Network Models of Memory*. PhD thesis, KTH, Sweden, 2003.

W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.

R. E. Suri and W. Schultz. Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Experimental Brain Research*, 1998.

R. E. Suri and W. Schultz. A neural network model with dopamine-like reinforcement signal that learns a spartial delayed response task. *Neuroscience*, 1999.

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction.* MIT Press, 1998.

H. C. Tuckwell. *Introduction to Theoretical Neurobiology.* Cambridge University Press, 1998.

J. Wickens. Basal ganglia : structure and computations. *Comput. Neural. Syst.*, 8: R77–R109, 1997.