Biol. Cybern. 73, 265–274 (1995)

Biological Cybernetics

Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning

Peter F. Dominey

Vision et Motricité, INSERM Unité 94, F-69500 Bron, France

Received: 21 July 1994/Accepted in revised form: 21 March 1995

Abstract. A novel neural network model is presented that learns by trial-and-error to reproduce complex sensory-motor sequences. One subnetwork, corresponding to the prefrontal cortex (PFC), is responsible for generating unique patterns of activity that represent the continuous state of sequence execution. A second subnetwork, corresponding to the striatum, associates these state-encoding patterns with the correct response at each point in the sequence execution. From a neuroscience perspective, the model is based on the known cortical and subcortical anatomy of the primate oculomotor system. From a theoretical perspective, the architecture is similar to that of a finite automaton in which outputs and state transitions are generated as a function of inputs and the current state. Simulation results for complex sequence reproduction and sequence discrimination are presented.

1 Introduction

Most intelligent behavior is performed in some context, making it important for the nervous system to maintain an internal representation of this behavioural context. From a systems perspective, we can call this representation 'state,' defined as a cumulative record of the previous inputs and outputs that is used to determine the response of the system to new inputs. In this fashion, all events become connected in a sequential fashion, via their contribution to the current state, and the current state's contribution to subsequent actions.

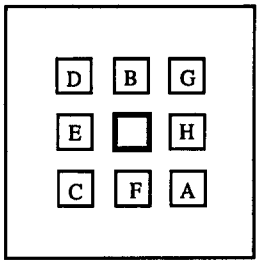
For sensory-motor systems, the term 'sequence reproduction' is misleading in the sense that such systems rarely produce sequences, but instead they perform translations from sensory input to motor output formats, as in hearing and then repeating a spoken phrase.

Here a task is considered in which a sequence of spatial targets is presented on a visual display. After a brief delay, the subject (model) must reproduce the sequence by orienting to the targets in the same order that they were presented. Thus the visually perceived sequence is translated into an equivalent motor output sequence. A related model (Dominey 1993; Dominey et al. 1995) was originally developed to describe the activity of prefrontal cortical (PFC) cells recorded in trained primates that performed an oculomotor version of this task (Barone and Joseph 1989). Dominey et al. (1995) focused on single unit activity in PFC for only the most simple (three elements) sequences. The current report looks specifically at the more difficult problem of learning complex sequences of up to 20 elements with repeated sub-sequences, and how the required encoding of sequence state takes place in the network corresponding to PFC. In addition, the model's capacity to both reproduce and discriminate between multiple sequences is examined, and different learning paradigms are compared.

In addressing the motor expression of a sensory input sequence, the model differs from many sequence learning models (e.g. Herz et al. 1989; Dehaene et al. 1987; Wang and Arbib 1990) that are designed to allow the storage and retrieval of sequences of imposed network states (see Kühn and van Hemmen 1992 for an extensive review.) In the current system, a sequence of network states generated by sensory input (and recurrent motor output signals) become associated by reinforcement learning with a corresponding sequence of states in the output network.

These temporal sequences can be characterized in terms of their length, complexity and order. The length is the number of elements in the sequence. The complexity refers to the maximum number of symbols that must be remembered in order to know the correct successor. Consider, for example, the sequence *A-B-A*-D-*A-B-A*-E-*A-B-A*-F. In order to correctly produce 'E', the system must remember the four previous elements that define the context for E; thus, the complexity of this sequence is four. The order is a measure of hierarchical complexity: if a recurring sub-sequence (e.g. *A-B-A* above) has within itself a recurring sub-sequence (e.g. A in *A-B-A*), the sequence is a high-order complex sequence, otherwise a first-order complex sequence (see Wang and Arbib 1990).

The sequencing paradigm employed here is an extended version of that introduced by Barone and Joseph (1989). A sequence $(s_1, s_2, \ldots s_n)$ of visual inputs on



Sensory input: A - B - C - D - g - g - g - g
Motor output: A'- B' - C' - D'

Fig. 1. Sequence reproduction paradigm. Visual input array of spatial targets labelled *A-H*. Input is a temporal sequence of activation of these spatial targets, e.g. *A-B-C-D*. After a delay, the task is to produce the motor output sequence of orienting movements to the spatial locations, denoted *A'-B'-C'-D'*, in the same order of the initial sequence presentation. Each output is triggered by a go-signal (*g*), which consists of phasic activation of all targets simultaneously

a 5×5 array is presented as shown in Fig. 1. After a delay, a sequence ($g_1, g_2, \ldots g_n$) of go signals is presented. Each go signal consists of presenting elements *A–H* simultaneously. For each go signal g_i, the *i*th item in the sequence must be chosen.

2 Complex sequence learning model

The model is based on a plausible biological implementation of the following functional requirements that partition the task of sequence reproduction into a problem of state representation, and a second problem of associative memory.

1. State representation: At any point during the presentation or reproduction of the sequence, the internal state of the system encodes the history of the sensory and motor events that have so far occurred.
2. Associative memory: At each point during the sequence reproduction, the current state will become linked, via reinforcement learning in a simple associative memory, to the correct next output in the sequence.

In the model, the state representation function is performed by a network corresponding to PFC. Barone and Joseph (1989) recorded single PFC units whose spatially and temporally selective activity encoded the ongoing state of sequence presentation and execution. The associative memory is implemented in a set of modifiable connections between PFC and the caudate nucleus. Kermadi et al. (1993) recorded single units in the caudate whose activity reflected a conditional association between sequence state and associated motor responses. The projection from prefrontal cortex to caudate (Selemon and Goldman-Rakic 1985) provides an anatomical basis for this associative memory.

Figure 2A displays a schematic representation of this system. Sensory and motor information influence the internal state, and an associative memory links internal states to motor outputs. Figure 2B shows the progression

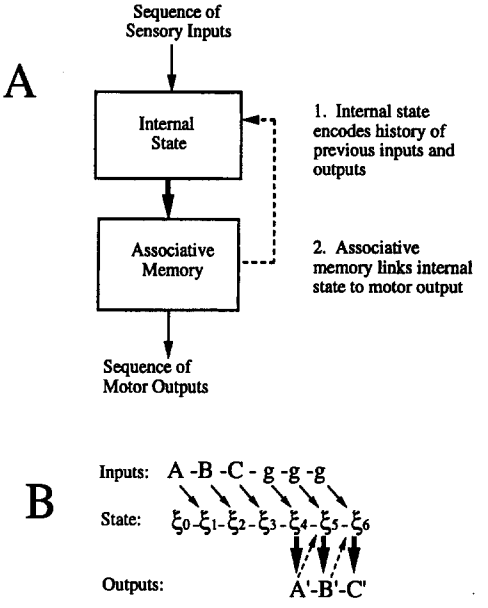


Fig. 2. Sequence reproduction schema. Internal state represents history of sensory inputs and motor outputs. State-response associative memory associates internal states with the correct next response in the sequence, via reinforcement learning. See text for details

in time of the reproduction of sequence *A-B-C*. The model starts in an initial state ξ_0. Presentation of the first visual input *A* drives the system to a new state, ξ_1. Presentation of input *B* drives the system from ξ_1 to ξ_2. Then presentation of input *C* drives the system from ξ_2 to ξ_3. Presentation of the first go signal (*g*) drives the model to state ξ_4 and triggers the model to produce a motor output by retrieving from the associative memory the output currently associated with state ξ_4 (dark vertical arrow in 2B).

If the retrieved output is incorrect, the offending association is weakened, reducing the probability that the same choice will be made again. If the output is correct, *A'*, this association is strengthened, and the system moves on to state ξ_5 which retrieves *B'* from the associative memory by the same process, and so on. By this trial and error learning, this system will learn the state-output associations (indicated by the heavy arrows in Fig. 2A and B) and thus will reproduce spatiotemporal sequences as the concatenation of state-response associations.

2.1 Model organization

The model is implemented in Neural Simulation Language (NSL; Weitzenfeld 1991), in which the base component is the two-dimensional layer of units, each corresponding to a population of neurons. The internal state of each layer is represented by a single variable $m(t)$, a 'leaky integrator' which is described by an array of differential equations of the form:

$$\tau_m \, dm(t)/dt = -m(t) + S_m(t) \qquad (0.1)$$

Here τ_m is the time constant for the rate of change of $m(t)$. $S_m(t)$ represents the total input that cells of type m receives from other cells. The Euler method is used to

solve the differential equation. The firing rates are determined by a non-linear function of m:

$$M = \text{sigmoid}(m, \text{min}, \text{max}, \text{MIN}, \text{MAX}) \quad (0.2)$$

where M is the firing rate, min and max are the minimum and maximums that define the input range for m, and MIN and MAX define the limits on the output firing rates. For min $< m <$ max, the firing rate is a non-linear function of m:

$$M = (\text{MAX} - \text{MIN})^* (S^{2*} (3 - 2S)) + \text{MIN}$$

where

$$S = (m - \text{min})/(\text{max} - \text{min}), \quad (0.3)$$

For $m <$ min or $m >$ max, the firing rate is MIN or MAX respectively. For each layer, the following parameters are specified: the time constant, the inputs to the layer, and the threshold values on the sigmoid function used to compute the firing rate from m. Values for these parameters are provided in the appendix. In general, these values are chosen to give the simulated neural elements the same connectivity and range of spatiotemporal activation (firing rate, latency etc.) as seen in their neuroanatomical counterpart as revealed by primate neuroanatomy and electrophysiology experiments.

The notation is simplified by defining a function $g()$ where $y = g(m)$ represents that the value of $m(t + \Delta t)$ is determined by calculating $m(t)$ and the subsequent firing rate as described above.

The sensory-motor model architecture (Fig. 3) is based on the cortical and subcortical organization of the oculomotor saccade system (Dominey and Arbib 1992). The input consists of a temporal sequence of visual targets in the 5×5 spatial array Visual_Input. The input is projected onto the retina as a function of the current eye position (EP).

$$\text{Retina}(i, j) = \text{Visual\_Input}(i + \text{EP}_{i,\,j} + \text{EP}_j) \quad (1)$$

After a delay, the model must reproduce the input target sequence as an output motor sequence of saccades to the targets in the order they were presented. Outputs are generated as orienting movements (i.e. saccadic eye movements) encoded by topographic activity in the superior colliculus (SC), a topographic map of saccade amplitude and direction (Sparks 1986). SC receives excitatory saccade command inputs from the frontal eye field (FEF) (Stanton et al. 1988b) and a tonic inhibition from substantia nigra (SNr; Chevalier et al. 1985). INH represents an inhibitory signal that is present during active fixation (Munoz and Wurtz 1992). The function $g'(x)$ is a 'winner-take-all', or 'maximum selector' version of $g(x)$ in which the maximum value in the array is retained and all others are set to zero.

$$\text{SC} = g'(\alpha\text{FEF} - \beta\text{INH} - \chi\text{SNr}) \quad (2)$$

Each SC-generated saccade centers the sensory input array (e.g. the retina) on the chosen target, via the recalculation of horizontal and vertical eye position (EP_i, EP_j, respectively). The function $h_i(\text{SC})$ returns the horizontal component index of the active unit in SC, and

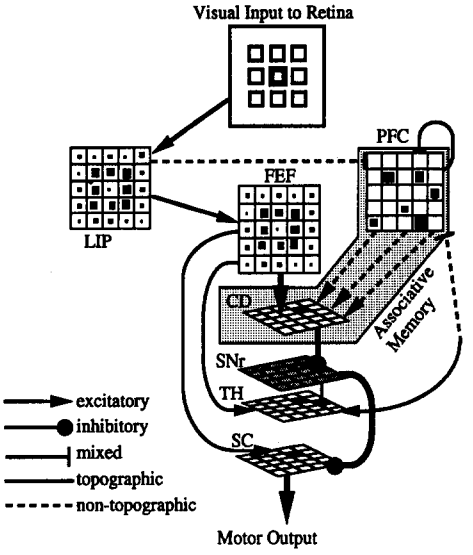


Fig. 3. Sequence model architecture. During saccade generation, continuous inhibition of superior colliculus (*SC*) by substantia nigra (*SNr*) is temporarily interrupted by caudate (*CD*) inputs. CD saccade-related cells are influenced by topographic projections from frontal eye fields (*FEF*) and also by modifiable, non-topographic projections from prefrontal cortex (*PFC*). PFC combines visual, motor (saccade efferent copy) and self input in order to generate a time-varying sequence of internal states for each presaccade period in the sequence reproduction task. These states or patterns of activity become associated with caudate activity for the correct saccade by reinforcement learning. *LIP*, lateral intraparietal area. See text

$h_j(\text{SC})$ returns the vertical component index. In (3) and (4) these displacement indices are combined with the previous eye position to yield the new eye position. The combination of (1)–(4) then serves as a high-level model of the oculomotor plant that repositions the eye (and thus the retina) based on saccade commands from SC.

$$\text{EP}_i(t + \Delta t) = h_i(\text{SC}(t)) + \text{EP}_i(t) \quad (3)$$

$$\text{EP}_j(t + \Delta t) = h_j(\text{SC}(t)) + \text{EP}_j(t) \quad (4)$$

The visual information in the retina then projects (via the thalamus and several intermediate cortical areas) to a layer corresponding to the saccade-related lateral intraparietal area (LIP) of the posterior parietal cortex in the primate (Van Essen 1979; Barash et al. 1991). Here, background noise [NOISE in (5)] is added to the visual input activity. The noise varies between 0 and 1, and is then scaled by β. Noise can provide a form of guessing that is useful for learning. LIP then projects this saccade-related information to the FEF (Andersen et al. 1985; Van Essen 1979).

$$\text{LIP} = g(\alpha\text{Retina} + \beta\text{NOISE}) \quad (5)$$

$$\text{FEF} = g(\alpha\text{LIP}) \quad (6)$$

To prevent movement during the presentation of the sequence, while a target is visually fixated (encoded by activation of the central element (0, 0) of LIP), element INH (7) is active, and is subsequently used as an inhibitory signal, as in (2). This inhibitory element is based on the fixation cells in FEF (Segraves and Goldberg 1987) that may be involved in suppressing movements during

active fixation (Munoz and Wurtz 1992). Removal of the FP with presentation of the remaining target(s) provides the go signal.

$$\mathrm{INH} = \mathrm{LIP}(0, 0) \tag{7}$$

Finally, the motor output activity is SC is inhibited by the substantia nigra (SNr), one of the main output nuclei of the basal ganglia, and SNr is itself inhibited by the caudate, one of the main input nuclei of the basal ganglia (Chevalier et al. 1985). The caudate in turn receives excitation from multiple cortical regions, including the FEF (Alexander et al. 1986; Selemon and Goldman-Rakic 1985; Yeterian and Pandya 1991). By this indirect path to SC via the basal ganglia, the cortex can then influence saccade generation (Hikosaka et al. 1989). Indeed, this indirect pathway will provide the basis for the associative memory (see Rolls and Williams 1987), described in Sect. 2.3.

$$\mathrm{SNr} = g(\alpha - \beta \mathrm{CD}) \tag{8}$$

$$\mathrm{CD} = g(\alpha \mathrm{FEF} - \beta \mathrm{INH}) \tag{9}$$

At this point a sensory-motor system that can produce saccades to spatial targets has been defined. The mechanisms for (a) representation of internal state, and (b) use of the state information to drive context-dependent movements will now be presented.

2.2 State representation

According to the requirement for state representation, the internal state mechanism should have access to information that defines the current sensory inputs, and the history of previous sensory and motor events. Equation (10) describes the dynamics of a 5×5 array of neurons, PFC, whose activity encodes the sequence state. Information related to the current sensory input is provided to PFC by LIP (Andersen et al. 1985; Goldman-Rakic 1987). W^{L-P} is a 25×25 matrix that defines this projection from LIP to PFC. The weights are fixed and vary between -0.55 and 0.45. This distribution of positive and negative weights (with a slight negative bias) provides excitatory and inhibitory influences in PFC, yielding a source of diverse activation of PFC that is not excessively excitatory, and thus will not drive PFC to saturation.

$$\mathrm{PFC} = g(\alpha \mathrm{LIP} * W^{L-P} + \beta \mathrm{SC_D} * W^{S-P} + \chi^{PF} \mathrm{C_D} * W^{P-P} + \delta \mathrm{TH}) \tag{10}$$

$$\mathrm{SC_D} = g(\alpha \mathrm{SC}(t)) \tag{11}$$

$$\mathrm{PFC_D} = g(\alpha \mathrm{PFC}(t)) \tag{12}$$

$$\mathrm{TH} = g(\alpha \mathrm{PFC} + \beta \mathrm{FEF} - \chi \mathrm{SNr}) \tag{13}$$

Information related to the most recent motor output is provided by SC_D, a damped (low-pass filtered) version of the SC output. In the primate it is likely that this information passes from colliculus to cortex indirectly via the thalamus (Schlag and Schlag-Rey 1984). W^{S-P} is a 25×25 matrix that defines the projections from SC_D to PFC. The weights are fixed and vary between -0.5 and 0.5. In order for previous states to influence current state representations in PFC, there are two recurrent pathways. PFC_D is a population of cells that get topographic (i.e. point to point) input from PFC and project back to PFC via the matrix W^{P-P} with fixed weights between -0.55 and 0.45, with all the self connections (i.e. W_{ii}^{P-P}) set to zero. The 25 PFC_D cells are subdivided into 5 groups of 5 cells that have a distribution of time constants (0.1, 0.6, 1.1, 1.6 and 2.1). This cortico-cortical connection between adjacent components of PFC (Goldman-Rakic 1987) provides PFC with a range of temporal sensitivity (see Sect. 3.2), similar to that provided by a distribution of temporal delays (Kühn and van Hemmen 1992). The second recurrent input is a topographic input from the thalamus layer TH, which receives PFC input. This is part of the prefrontal loop connecting cortex, basal ganglia and thalamus (Alexander et al. 1986). The visual, motor and recurrent information provide PFC with the required components for correct state maintenance, as illustrated schematically in Fig. 2B. The manifestation of this state encoding as diverse patterns of activity in PFC during sequence reproduction will be demonstrated in Sect. 3. Indeed, Barone and Joseph (1989) recorded single units in PFC while monkeys repeated simple oculomotor sequences and found that individual cells encoded a combination of spatial and temporal aspects of the targets in the sequence. The combined population would then presumably encode all the relevant state information required for the sequence execution.

2.3 Associative memory

An associative memory allows the activity in PFC encoding sequence state to command the correct saccades corresponding to each element of the sequence. The modeled substrate of this associative memory is in modifiable synapses connecting PFC (state) to CD (motor influence). Neurophysiological and anatomical evidence suggest that the cortico-striatal system may provide a basis for this associative memory (Rolls and Williams 1987). This has been substantiated by experiments demonstrating that the caudate is required for sensory-motor association learning (Reading et al. 1991; Robins et al. 1990). This plasticity may be based on dopamine-related modification of corticostriatal synapses after rewards (or absence of predicted rewards) since (a) there is a phasic modulation of dopamine release in the striatum with reward (or in the absence of a predicted reward) (Ljungberg et al. 1991), and (b) dopamine participates in synaptic plasticity in the striatum (Calabresi et al. 1992). Equation (9) is thus updated to include this adaptive projection from PFC to CD, implemented as this set of modifiable connections, W^{P-C}. The synapses in W^{P-C} are modified according to a reinforcement learning rule (see Barto 1990) after each saccade in a sequence execution (14). That is, for each choice in the sequence execution, synapses are strengthened between PFC cells encoding the state and the CD cells encoding the response after correct responses, or weakened after

incorrect responses.

$$CD = g''(\alpha FEF - \beta INH + \chi PFC^{*}W^{P-C}) \quad (9')$$

The standard value of R is typically 2.5e-5 for correct saccades, and -2.5e-5 for incorrect saccades, though these values can be manipulated to produce different learning strategies (see Sect. 3.4). To prevent saturation and provide a form of competition between patterns in the associative memory, the weights are normalized so that the total output projection strength for any cell in PFC is conserved. This algorithm is represented in (15) where the final value for W_{ij}^{P-C} is determined by taking the new value returned from (14) and scaling it by the relative change in the sum of synaptic weights issuing from unit PFC_i since the last update, so that this sum is always preserved. The parameter specific performance characteristics of this associative memory are discussed in Sect. 3.4.

$$W_{ij}^{P-C}(t+1) := W_{ij}^{P-C}(t) + R^{*}PFC_i^{*}CD_j \quad (14)$$

$$W_{ij}^{P-C}(t+1) := W_{ij}^{P-C}(t+1)^{*}\frac{\sum_j W_{ij}^{P-C}(t)}{\sum_j W_{ij}^{P-C}(t+1)} \quad (15)$$

During and after learning, it is possible that the combined inputs to CD will exceed the upper boundary (max) of the non-linear functions, thus saturating this function. If this occurs such that both correct and incorrect choice representation in CD exceed the threshold, the information necessary to make the correct choice will be lost in the saturation. To accommodate this problem, we perform a normalization in $g''(x)$ such that if the maximum value in array x (x.max) > the input upper limit (max), we divide all elements in array x by the term 'max/x.max'. This ensures that the threshold function never saturates, and thus that information regarding the relative strength of different elements in the array is never lost to saturation. This modulatory function appears to be carried out in the striatum by modulation of striatal dopamine levels which regulates corticostriatal excitability (Mercuri et al. 1985) in a negative feedback loop (Carlsson and Carlsson 1990).

3 Simulation results

The learning capacities of this network for single and multiple complex sequences that have length of up to 20 items, and repeated sub-sequences are investigated in detail. Consider the sequence *A-B-C*-D-*A-B-C*-E-*A-B-C*-F-*A-B-C*-G-*A-B-C*-H. Here the sub-sequence *A-B-C* is repeated five times, each time followed by a different response – D, E, F, G and H, respectively. In order to learn and reproduce this sequence, the state mechanism must be capable of resolving the redundancy when *A-B-C* is observed in order to choose the correct successor. This requires a functional memory span of 4 (i.e. equal to the complexity of the sequence). Simulation results will show how the state mechanism demonstrates this memory span, and how the associative memory uses this state information to retrieve the corresponding motor responses.

3.1 Complex sequence reproducing

The training paradigm resembles those used in animal training, where the desired behavior is shaped gradually. Training was divided into 4 epochs. In the first "instructed" epoch, the 20 successive targets were shown during the sequence presentation, and then during the reproduction, instead of presenting the standard go signal (illumination of all the targets) 20 successive times, only the correct target was presented at each of the 20 response steps. That is, there was no choice involved and thus no errors. The model simply picked the only available target at each step. The learning rate was set to 1e-4 during this period. Each of the three remaining training epochs (2–4) were divided into 11 phases. Each phase was repeated until completed successfully 4 times before moving on to the next phase. Phase 1 was 'instructed' as defined above. In phase 2, the first 2 go signals were 'non-instructed' – they consisted of presentation of all the targets, so the model had to choose correctly for the first two items in the sequence. In phase 3, the first 4 go signals consisted of presentation of all the targets, so the model had to choose correctly for the first 4 items in the sequence.

In each successive phase, the subsequent two targets had to be successfully chosen from all the targets. In phase 11, the entire sequence was performed with all choices made from all of the 8 targets. The probability against this is $1:8\times7^{19}$ (i.e. 8 choices for the first item, then 7 for each of the remaining choices). Once the 11 phases had been successfully repeated 4 times each (i.e. once an epoch was complete), the entire process was then repeated again. After 4 such training epochs, the sequence was performed perfectly. Table 1 describes the error rates and learning parameters for the different epochs.

Figure 4A illustrates the activity of the 25 PFC cells during the presentation and reproduction of the complex sequence *A-B-C*-D-*A-B-C*-E-*A-B-C*-F-*A-B-C*-G-*A-B-C*-H,

Table 1. Performance parameters for sequence: *A-B-C*-D-*A-B-C*-E-*A-B-C*-F-*A-B-C*-G-*A-B-C*-H. This is a complex sequence of degree 4

	Total	Number correct	Number incorrect	Percent correct	Learning rate	Forgetting rate
Epoch 1	48	48	0	100	1e-4	0
Epoch 2	135	65	70	48	5e-6	2.5e-5
Epoch 3	124	107	14	86	1e-6	1.5e-5
Epoch 4	115	113	2	98	0	1.5e-5

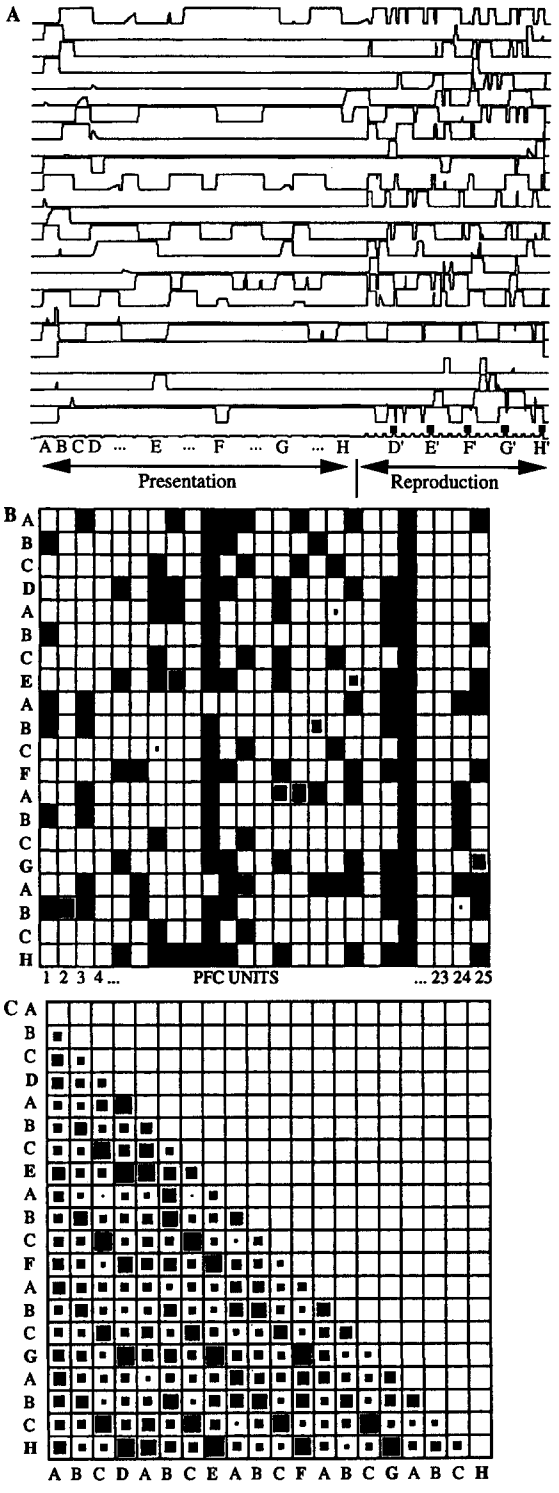


Fig. 4. **A** State-encoding activity of 25 PFC cells during execution of sequence A-B-C-D-A-B-C-E-A-B-C-F-A-B-C-H. Time runs from *left* to *right*. *Bottom trace* indicates sequence progress during presentation of 20 targets and 20 go-signals and the corresponding orienting movements. *Squares* indicate go-signals for the redundant movements *D'*, *E'*, *F'*, *G'* and *H'*. **B** Each *horizontal row* represents the level of activation (range 0 to 100) of the 25 PFC cells at the time that the output (indicated at the *left*) was generated. These are the patterns in PFC that encode state, and become associated with activation of CD for the correct saccade. **C** Each element (range 0 to 1) represents the cosine of the angle between the two state vectors indicated on the vertical and horizontal axes (see text)

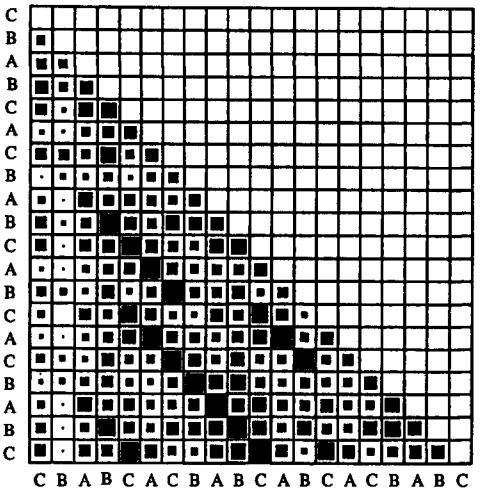


Fig. 5. Same format as Fig. 4C, for sequence C-B-A-B-C-A-C-B-A-B-C-A-B-C-A-C-B-A-B-C

with time running from left to right. The periods of target presentation and sequence reproduction are indicated on the bottom trace. During the visual presentation and motor orientation to each target (indicated on the last trace), there is an evolving pattern of activity in the population of PFC cells. While the sequence of motor responses is complex (i.e. the sub-sequence *A-B-C* is repeated 5 times, each time with a different successor), the sequence of patterns in PFC is simple, in the sense that none of the activity patterns are repeated. In order to visualize the relations between the different states encoded in PFC during the reproduction, a set of 25 element vectors encoding the PFC activation at each output point in the sequence reproduction is extracted (Fig. 4B). A measure of similarity (the cosine of the angle between each pair of these two vectors) is calculated in order to display the degree of overlap between these patterns. As the vectors become more similar, this value approaches 1.

In Fig. 4C, this matrix of similarity measured is displayed between all pairs of PFC vectors during execution of the complex sequence in Fig. 4A. The repetitive structure of this sequence can be seen as diagonal bands that indicate high correlations between certain states. On the last row (comparing state H' to all other states) we see that the 'redundant' states D', E', F', and G' are all fairly similar to state H', since for each a previous sub-sequence of outputs (A'-B'-C') has been the same. Following the diagonals left- and upward, we can see that the predecessor states are also similar. While some states are similar, they are sufficiently distinct that the associative memory can distinguish them and recall the correct output for resolving the redundancies.

Figure 5 is similar to that of Fig. 4C, but for the sequence C-B-A-B-C-A-C-B-A-B-C-A-B-C-A-C-B-A-B-C. Again note in Fig. 5, like Fig. 4C, how the structure of the sequence is revealed as diagonal bands in the correlation matrix. This is a 'high order' complex sequence, because the repeated sub-sequence C-B-A-B-C-A is itself complex (i.e. C and B both have non-unique successors). Table 2 presents the learning parameters for this

Table 2. Performance parameters for sequence: C-B-A-B-C-A-C-B-A-B-C-A-B-C-A-C-B-A-B-C

	Total	Number correct	Number incorrect	Percent correct	Learning rate	Forgetting rate
Epoch 1	6	6	0	100	5e-5	2.5e-5
Epoch 2	62	55	7	88	1e-5	2.5e-5
Epoch 3	60	60	0	100	1e-6	1.5e-5

Table 3. Learning parameters for sequence discrimination

	Total	Number correct	Number incorrect	Percent correct	Learning rate	Forgetting rate
Epoch 1	241	237	4	98	1e-4	0
Epoch 2	596	447	149	75	5e-6	2.5e-5
Epoch 3	601	592	9	98	1e-6	1.5e-5
Epoch 4	601	601	0	100	0	1.5e-5

sequence. Note that this sequence is learned more efficiently than that described in Table 1, due in part to the fact that in this sequence there are only 3 targets, vs. 8 in the previous sequence, with probabilities $1:3 \times 2^{19}$ versus $1:8 \times 7^{19}$, respectively.

3.2 Variations on the standard protocol

The model is relatively flexible with respect to the number of sequences that can be learned in given training session, and changes in the timing of targets and go signals. As the sequences become increasingly complex, however, this flexibility is reduced. Thus, in a single training session, the model is capable of learning multiple sequences, dependant on their length and complexity, e.g. 3 simple sequences of length 20; 6 simple sequences of length 3; versus 2 complex sequences of length 12.

Likewise, the timing delay between a motor response and the subsequent go signal can be increased up to 200% with no perturbation for simple sequences, while an increase of 50% impairs performance for complex sequences. The distribution of time constants in PFC_D contributes to this relative flexibility. Indeed, the range of these time constants can be modified from 0.1–2.1, to 0.01–1.01 still leaving performance intact. This is in agreement with the observation of Herz et al. (1989) that with a distribution of internal delays, the network capacity to encode spatio-temporal information is robust, and largely independent of the details of the distribution. That this flexibility is reduced for complex sequences, which place a greater load on the state and associative memory mechanisms is not surprising, as Stadler (1995) has observed similar perturbations in human complex sequence learning induced by timing modifications.

While the standard protocol described above uses the initial presentation of the sequence as an index for its subsequent retrieval, the model can also operate immediately in retrieval mode where presentation of the initial 2 or 3 targets during sequence reproduction acts as an index for the subsequent retrieval of the rest of the sequence. Thus the model was trained to error free performance on three different sequences of length 20, with no initial presentation of the sequence. Instead, the first three elements of each sequence, presented in the 'instructed' mode, served as the index for retrieval of the rest of the sequence in the 'non-instructed' mode, similar to the method employed by Dehaene et al. (1987). These variations on the standard protocol illustrate the flexibility of this architecture.

3.3 Sequence discrimination

Another interesting and straight-forward result of this architecture is that it can be used for sequence recognition/discrimination. After a sequence is presented, PFC will be in a given state. By trial and error the model learns to associate this state with a given response and to discriminate it from the state induced by other sequences. This capability was tested by using four sequences: (1) A-B-C-D-E-, (2) A-B-C-E-D, (3) D-B-A-C-E and (4) E-C-A-B-D, and arbitrarily assigning to each one a motor response (saccades to locations G, F, C and A, respectively). Table 3 provides the training and performance details for learning to recognize and discriminate these four sequences.

An interesting observation concerning the training schedule adopted for this task is presented in the following section.

3.4 Learning strategies: training format, learning and forgetting

In this model, correct behaviour can be characterized in terms of learning the associations between patterns of activity in PFC, and activation of units in CD that correspond to the associated response. Here we consider three parameters that can be manipulated to improve the learning performance: (1) learning rate, (2) forgetting rate and (3) the training schedule.

For complex sequences a high learning rate during was set in the first phase (i.e. the instructed phase in which the no choices are required) to provide a good first approximation to the desired behavior. When the choices

are introduced, some errors are made, and the early elements of the sequence are repeated more than later elements, since the protocol always re-starts at the beginning of the sequence after the first error is made. By the competition of weight normalization (15), the associations for the latter elements are successively weakened with respect to the early elements that become over-learned. In this configuration the system demonstrates a tendency to get stuck, never progressing beyond a certain point in the sequence. To address this problem, an initially high learning rate during the instructed period was used, as before, and then when choice was introduced the learning rate was reduced and the forgetting rate increased so that synaptic changes occurred only when mistakes were made (see Tables 1 and 2). This eliminates the destructive competition from over-learning early components, and allows convergence to the correct behavior. This is a simple demonstration that the desired performance can be attained by a combination of increasing the probabilities for correct behavior and decreasing the probabilities for incorrect behavior by 'learning' and 'forgetting,' respectively.

The importance of training schedule was revealed in the sequence discrimination experiments. The initial schedule was organized similar to that for sequence reproduction learning. In an instructed period, each sequence is presented, and then only the correct choice is presented, and the model responds correctly. In the non-instructed period, the sequence is presented and the model chooses from four responses. If a correct response is made, the protocol moves on to the next sequence. If an incorrect response is made, the sequence is repeated until the correct choice is made. An instability resulting from competition between two sequences then developed. Sequence 3 repeated several times until it correct, then when sequence 1 was later presented, it repeated several times until correct in a cyclic behavior. That is, the changes for learning 3 interfered with what had been learned for 1, then the reverse occurred in an unstable cycle. The result of this failure is seen in Table 4. In order to reduce this interference, the schedule was changed so that after an error is made [and the synaptic changes by (14) and (15) are made] the protocol returns to the beginning of the phase, i.e. to sequence 1 (rather than to the beginning of the sequence in which the error was made). In this manner, each time the associative memory is modified, all of the previously learned associations are updated before proceeding. The difference can be seen in comparing Tables 3 and 4.

4 Discussion

A new sequence learning model is presented that is robust both in the number of sequences it can learn in a given training session (e.g. two sequences, each of length 12, complexity four; or 3 simple sequences of length 20 stored in the same set of synapses) and in the length and complexity (20 elements, complexity 4) of sequences it can learn, placing it in good competition with existing models (e.g. Dehaene et al. 1987; Herz et al. 1989; Wang and Arbib 1990). The novelty of this sequence learning model is that it employs an architecture that decomposes the sequencing task into two components, based on a general corticostriatal organization of the primate brain (Alexander et al. 1986), as typified in the oculomotor system (see Dominey and Arbib 1992; Dominey et al. 1995). One sub-network, corresponding to the primate prefrontal cortex, is responsible for generating unique patterns of activity in a group of neurons, where the patterns represent the current state of the sequence execution. A second sub-network, corresponding to the primate striatum, associates these state-encoding patterns with the correct response at that point in the sequence execution. Correct guesses strengthen the connections between the active state (PFC) neurons and the active motor (caudate) neurons forming an associative memory by reinforcement learning.

This reduction of sequence processing into state representation and associative memory illustrates an important point, as it allows the possibility of a separate state mechanism that can interact independently with multiple output systems. That is, a sequence may be performed by looking at the targets, or pointing to them with the left or right hand. In all three cases, there is a common sequencing element, and it seems more efficient to exploit this commonality in a separable function, rather than duplicating it for each output mode. This kind of transfer between modular effector systems has been demonstrated for sequence learning in humans (Keele et al. 1995). Anatomically, this modularity is reflected by the fact that the cortex – basal ganglia loops are not completely segregated but have some overlap allowing one system to influence the other(s) (Alexander et al. 1986). It is highly likely that there are other brain mechanisms for sequence learning in addition to the proposed corticostriatal plasticity mechanism.

Among the greatest challenges for this kind of sequence learning system are those imposed by complex sequences, since the degree of complexity imposes additional requirements on the memory span of the sequence.

Table 4. Parameters for an inefficient sequence discrimination training schedule

	Total	Number correct	Number incorrect	Percent correct	Learning rate	Forgetting rate
Epoch 1	241	237	4	98	1e-4	0
Epoch 2	591	32	559	5	5e-6	2.5e-5
Epoch 3	590	5	585	1	1e-6	1.5e-5
Epoch 4	590	5	585	1	0	1.5e-5

A simple sequence can be learned as a set of associations between patterns ξ^{ν} and $\xi^{\nu+1}$, since each $\xi^{\nu+1}$ is uniquely determined by ξ^{ν}. In complex sequences this is not the case, a given ξ^{ν} may have multiple successors. As pointed out by Kühn and van Hemmen (1992), there are at least two ways to approach this constraint. First, the system may be endowed with 'time delays of sufficient length to memorize as much of the history of the network evolution as is necessary to resolve any ambiguity that might come along as the complex sequence is traced out.' (p. 261.) The second method involves structuring the network, as that of Dehaene et al. (1987), so that even though the activation of the output clusters describes a complex sequence, taking into account the activation of the internal clusters as well, the network goes through linear sequences.

In this sense, the model presented in this report is similar to that of Dehaene et al. (1987), in that the sequence of patterns produced in PFC is a simple sequence (i.e. no repeated sub-sequences), while the pattern of motor outputs is complex. This is achieved by associating each of the states generated in PFC with its corresponding motor output.

Because this model learns to reproduce and recognize sequences in a trail-and-error fashion based on known brain organization, its behavioural results can eventually be used to make predictions for normal and diseased human performance in neuropsychological tests.

Acknowledgements. This work was supported by a postdoctoral fellowship to P. D. from the Direction de la Recherche et des Etudes Doctorales (DRED), a grant from the Human Frontiers Organization, and by Vision et Motricité, INSERM Unité 94, directed by Professor Marc Jeannerod. The author also acknowledges the insightful and useful comments of an anonymous reviewer.

Appendix: Equation parameters

	t_m	min	max	MIN	MAX	α	β	χ, δ
(2)	0.008	60	100	0	100	2	2	1
(5)	0.01	0	75	0	110	70	5	–
(6)	0.01	0	100	0	100	1	–	–
(8)	0.01	0	75	0	100	1	–	–
(9)	0.01	40	100	0	75	0.8	0.4	0.15
(10)	0.01	40	110	0	100	5	8	2, 6
(11)	0.2	0	100	0	100	19	–	–
(12)	*[a]	0	100	0	100	*[b]	–	–
(13)	0.01	0	100	0	70	1.5	1	1

[a] The population of 25 cells is subdivided into five groups of 5, with respective time constants: .1, .6, 1.1, 1.6, 2.1.
[b] The constants, α, for these sub-populations are 1, 1.5, 2, 4, 6, respectively.

References

Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357–381

Andersen RA, Asanuma C, Cowan MW (1985) Callosal and prefrontal associational projecting cell populations in area 7a of the macaque monkey: a study using retrogradely transported flourescent dyes. J Comp Neurol 232:443–455

Barash S, Bracewell RM, Fogassi L, Gnadt JW, Andersen RA (1991) Saccade-related activity in the lateral intraparietal area II. Spatial properties. J Neurophysiol 66:1109–1124

Barone P, Joseph J-P (1989) Prefrontal cortex and spatial sequencing in macaque moneky. Exp Brain Res 78:447–464

Barto AG (1990) Connectionist learning for control: an overview. In: Miller WT, Sutton RS, Werbos PJ (eds) Neural Networks for Control. MIT Press, Cambridge MA, pp 1–58

Calabresi P, Maj R, Pisani A, Mercuir NB, Bernardi G (1992) Long-term synaptic depression in the striatum: physiological and pharmacological characterization. J Neurosci 12:4224–4233

Carlsson M, Carlsson A (1990) Interactions between glutamatergic and monoaminergic systems within the basal ganglia – implcations for schizophrenia and Parkinson's disease. Trends Neurosci 13:272–276

Chevalier G, Vacher S, Deniau JM, Desban M (1985) Disinhibition as a basic process in the expression of striatal functions. I. The striato-nigral influence on the tecto-spinal tecto-diencephalic neurons. Brain Res 334:215–226

Dehaene S, Changeux J-P, Nadal J-P (1987) Neural networks that learn temporal sequences by preselection. Proc Natl Acad Sci 84:2727–2731

Dominey PF (1993) Models of spatial accuracy and conditional behavior in oculomotor control. PhD Thesis, Dept of Computer Science, University of Southern California

Dominey PF, Arbib MA (1992) A cortico-subcortical model for generation of spatially accurate sequential saccades. Cerebral Cortex 2:153–175

Dominey PF, Arbib MA, Joseph JP (1995) A model of corticostriatal plasticity for learning oculomotor associations and sequences. J Cog Neuroscience 7:311–336

Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: Vernon B Mountcastle (ed) Handbook of physiology. The Nervous system. Vol. V, Chap 9. 9:373–417

Herz A, Sulzer B, Kühn R, van Hemmen JL (1989) Hebbian learning reconsidered: representation of static and dynamic objects in associative neural nets. Biol Cybern 60:457–467

Hikosaka O, Sakamoto M, Usui S (1989) Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. J Neurophysiol 61:780–798

Keele SW, Jennings P, Jones S, Caulton D, Cohen A (1995) On the modularity of sequence representation. J Motor Behavior 27: 17–30

Kermadi I, Jurquet Y, Arzi M, Joseph J-P (1993) Neural activity in the caudate nucleus of monkeys during spatial sequencing. Exp Brain Res 94:352–356

Kühn R, van Hemmen JL (1992) Temporal association. In: Domanay E, Hemmen van JL, Schulten K (eds) Physics of neural networks. Springer, Berlin Heidelberg New York, pp 213–280

Ljungberg T, Apicella P, Schultz W (1991) Responses of monkey midbrain dopamine neurons during delayed alternation performance. Brain Res 567:337–341

Mercuri M, Bernardi G, Calabresi P, Cotugno A, Levi G, Stanzione P (1985) Dopamine decreases cell excitability in rat striatal neurons by pre- and postsynaptic mechanisms. Brain Res 385: 110–121

Munoz DP, Wurtz RH (1992) Role of the rostral superior colliculus in active visual fixation on execution of express saccades. J Neurophysiol 67:1000–1002

Reading PJ, Dunnet SB, Robins TW (1991) Dissociable roles of the ventral, medial and lateral striatum on the acquisition and performance of a complex visual stimulus-response habit. Behav Brain Res 45:147–161

Robins TW, Giardini V, Jones GH, Reading PJ, Sahakian BJ (1990) Effects of dopamine depletion from the caudate-putamen and nucleus accumbes septi on the acquisition and performance of a conditional discrimination task. Behav Brain Res 38:243–261

Rolls ET, Williams GV (1987) Sensory and movement-related neuronal activity in different regions of the primate striatum. In: Schneider JS, Lidsky TI (eds) Basal ganglia and motor behavior. Hans Huner, Toronto, pp 37–59

Schlag J, Schlag-Rey M (1984) Visuomotor functions of central thalamus in monkey. II. Unit activity related to visual events, targeting, and fixation. J Neurophysiol 51:1175–1195

Segraves M, Goldberg ME (1987) Functional properties of corticotectal neurons in the monkey's frontal eye field. J Neurophysiol 58: 1387–1419

Selemon LD, Goldman-Rakic PS (1985) Longitudinal topography and interdigitation of corticostriatal projections in the rhesus moneky. J Neurosci 5:776–794

Sparks DL (1986) Translation of sensory signals into commands for control of saccadic eye movements: role of primate superior colliculus. Physiol Rev 66:118–171

Stadler MA (1995) The role of attention in implicit learning. J Exp Psychology: Learning, Mem, Cognition (in press)

Stanton GB, Goldberg ME, Bruce CJ (1988a) Frontal eye field efferents in the macaque monkey. I. Subcortical pathways and topography of striatal and thalamic terminal fields. J Comp Neurol 271: 473–492

Stanton GB, Goldberg ME, Bruce CJ (1988b) Frontal eye field efferents in the macaque monkey. II. Topography of terminal fields in midbrain and pons. J Comp Neurol 271:493–506

Van Essen DC (1979) Visual areas of the mammalian cerebral cortex. Annu Rev Neurosci 2:227–263

Wang D, Arbib M (1990) Complex temporal sequence learning based on short-term memory. Proc IEEE 78:1536–1543

Weitzenfield A (1991) NSL neural simulation language version 2.1. (Technical Report 91-05). Center for Neural Engineering, University of Southern California

Yeterian EH, Pandya D (1991) Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. J Comp Neurol 312:43–67