Chapter 6

PROPOSAL OF FUNCTIONS OF THE BASAL GANGLIA LOOPS

6.1 Introduction

Based on the reviews and discussions of the preceding chapters, the aim of this chapter is to propose the functions for the basal ganglia-thalamocortical loops in sequential movements from a computational viewpoint. First, the computational elements are discussed, and, second, a general framework of the functions of the loops is proposed in terms of the acquisition and retrieval processes in execution. These descriptions are given in relation to the 2x5 task. Third, a model implementing an algorithm based on the above framework is provided in order to perform a close comparison of the model with the experimental data of the 2x5 task, which is in the next chapter.

6.2 Computational elements in the basal ganglia thalamocortical loops for sequential decision making

For the following discussions, it may be worth introducing the distinction between the two kinds of memory of sequences: one depends on the sensory inputs (the *sensory-sequence* memory) and the other depends on the motor outputs (the *motor-sequence* memory). In the motor-sequence memory, the memory should really include information represented by the kinematical/dynamical representation at least to some degree. In visuo-motor task such as the 2x5 task, the motor-sequence memory includes nonlinear mappings, or roughly speaking, complex mappings. In contrast, the sensory-sequence memory can be represented in terms of the Cartesian representations, in particular, as a sequence depending on the spatial representation. For the saccade by the oculomotor system, the Cartesian representation of the target in the external world can be employed.

Note that, from a computational viewpoint, there generally exists a tradeoff between execution and acquisition in terms of the type of information stored in the memory. It is easier and faster to produce the output, or movements, if the information is stored more closely to the output, or downstream. At the same time, it is often more difficult to acquire the information close to the actual output but rather easier to acquire the information somewhat at an abstract level or upstream. Consider the case of reaching, for example. The Cartesian representation can be used to indicate an object as a target of reaching with respect to the position of a subject. The problem of the inverse-kinematics and inverse-dynamics should be solved with respect to the Cartesian representation of the target and the subject, to actually move a part of the body such as a hand to reach the target. It would be easier to acquire information for the sequences of reaching by Cartesian representation particularly in the early stage of acquisition, whereas it would be faster and easier to execute the sequences of reaching by the kinematical/dynamical representations once the memory of the sequence is acquired.

It can be said, thus, in the 2x5 task that the acquisition of the sequence may be much easier in terms of Cartesian representation than the kinematical/dynamical representation because the latter representation often becomes highly nonlinear to represent such sequences as employed in the 2x5 task. At the same time, it should be faster to perform the 2x5 task if the stored information in the memory is in the kinematical/dynamical representation rather than in Cartesian representation. Thus, it is plausible to consider from a computational viewpoint that the sensory-sequence memory is more suitable in the acquisition process, whereas the motor-sequence memory is more suitable in the execution process.

In the case of 2x10 task in the 'color' condition discussed in Section 5.4.2, the memory of color sequences is one of the sensory-sequence memory. In this case, the memory is independent of the spatial information. Though it is likely that the memory depends on the modality of colors, it may be possible for the memory to be stored at a somewhat abstract level. It is out of scope of the thesis to investigate whether a memory of the sequence exists at an entirely abstract level, which is completely independent of any sensory modality. Generally speaking, however, it is presumed that the sensory-sequence memory can be somewhat more abstract than the motor-sequence memory.

The distinction between the sensory-sequence and motor-sequence memories, however, should not be taken to imply that the sensory-sequence memory and the motor-sequence memory can be regarded as completely independent but rather as a conceptually valuable means for the investigation of the nature of the memory of sequences.

It should be also mentioned that, from a computational viewpoint, the term, 'context', is somewhat abused in some experimental literatures we have reviewed as well as in some computational models of the functions of the basal ganglia and related cortical areas. Roughly speaking, the term 'context' is used to refer to the situation in which the choice of an action taken at a state should be determined with respect to some of the events that happened in the past. In a series of experiments done by Schultz and his colleagues, discussed in Section 3.8, the term 'context' is used sometimes to indicate the procedure the monkeys engaged in and sometimes to indicate the experimental settings. Houk et al. [31] argued that neurons in the striatum are able to learn the contextual decision, positing functions of those neurons in the actor-critic scheme of reinforcement learning. The actor-critic scheme is able to take both of the immediate and long-term consequences into account and, consequently, is able to learn sequences with respect to its value function. This is a virtue of using reinforcement learning in sequential decision making. However, the actor-critic scheme by itself cannot be directly applied to learn a multiple of sequences, each of which is occasionally presented to the system. As a simple example, when there are two sequences each of which contains a particular state but each of whose action to be taken on the state is different, the actor-critic scheme will simply confuse both sequences. This kind of situation is often the case when people acquire skills in daily life as discussed in Section 3.1. It is so in the 2x5 task. It could happen that a correct order of actions, that is, pushing two illuminating buttons in a correct order, is different in the same set (the same sensory inputs) which belongs to different hypersets. From a computational viewpoint, thus, the case should be taken into account in which there are potentially multiple contexts, when the concept of context is considered in relation to sequential decision making. The context, or the sequence the subjects/monkeys are engaging, should be identified concurrently when they advance in the context, or at least be available for examination.

6.2.1 The acquisition processes in the early stage

When the monkeys encounter a new hyperset, they must explore what actions are correct for the sets of the new hyperset. In other words, the mapping of the sensory input (set) with the motor output (action) must be explored in the early stage of the acquisition. As seen in Section 5.4.2, the pre-SMA is known to be involved in the acquisition process in the early stage. The neurophysiological findings in the 2x5 task also clearly showed that the pre-SMA and the anterior striatum are much involved in the acquisition process in the early stage, discussed in Section 5.4. Based on discussion in Section 5.4.2, it is postulated in this study that the pre-SMA is responsible for the mapping of the sensory inputs to the motor outputs (the sensory-motor mapping), which should be particularly important in the early stage, in interaction of the anterior striatum (See also Section 5.4.1). Reinforcement signals are supposed to be provided by dopamine neurons in the substantia nigra pars compacta(SNc). Presumably, the monkeys would rather react to the sensory input to produce the motor output particularly in the very early stage of the acquisition, that is, at the beginning of encountering new hypersets. To produce the motor output, information for an action to the current sensory input in the pre-SMA should be transformed to be used. It is hypothesized in this study that the SMA transforms inputs from the pre-SMA to produce the motor output in ongoing process of making sequential movements in the early stage of the acquisition. As the monkeys get familiar with the hypersets via such exploration and have more successes, learning the sequence of the sets and their actions begins.

As reviewed in Section 3.6, the anterior striatum is involved in the dorsolateral prefrontal circuit and is much related to the oculomotor circuit. The pre-SMA is also known to be in close interaction with the dorsolateral prefrontal circuit and to have the projection to the anterior striatum. Neurons in the dorsolateral prefrontal cortex (DLPF) are known to be selective to a variety of characteristics of sensory inputs, particularly including spatial information [17, 18], and to have sustained activities, which can be considered as working memory. More importantly, the DLPF can be involved in determining sequential movements, depending upon the sensory information, as discussed in Section 3.6. In addition, as discussed in Section 5.4.2, the learning in the pre-SMA can be considered as more related to the sensory inputs rather than the motor outputs. Besides this evidence, the computational consideration of the tradeoff between the acquisition and execution discussed in the beginning of this section leads us to postulate that learning of the sequence in the early stage relies on the sequence of sensory inputs (the sensory-sequence memory) rather than that of motor outputs and that this learning occurs in the dorsolateral prefrontal circuit. In other words, the dorsolateral prefrontal circuit is hypothesized to learn the sensory-sequence memory given reinforcement signals by dopamine (DA) neurons in SNc.

Information of the sensory-sequence memory can be passed to the oculomotor circuit, for example, in the interaction of the DLPF and the supplementary eye field (SEF), to facilitate anticipatory saccades. In addition, information of the sensory-sequence memory in the DLPF can contribute to the production of the motor output in the SMA through the pre-SMA.

6.2.2 The acquisition process in the late stage

In contrast to the acquisition of the mapping of the sensory inputs to the motor outputs and the acquisition of the sensory-sequence memory, the nature of information used in the late stage is considered as depending more on the motor outputs. This view is supported by the experimental evidence such as (1) the learning measured by the number of errors to criterion decreases relatively faster than that measured by the performance time and (2) even after the 6 months interruption, by the performance time measure, the performance of learned hypersets is significantly better than that of new hypersets, whereas there is no significant difference in the number of error trials to criterion. In addition to the above evidence, information used in the late stage is rather sequential. This view is also supported by the experimental evidence that the performance for reversed hypersets was nearly as same as that for new hypersets, as discussed in Section 5.3. It is, thus, postulated that stored information in the late stage is the memory of the motor-dependent sequences (the motor-sequence memory). The SMA and the posterior putamen (Pt) in the motor circuit are considered as responsible for the acquisition of the motor-sequence memory, with reinforcement signals provided by DA neurons in SNc. As pointed out in Section 5.4.3, it may happen that the stored information in this circuit may be further transferred to other cortical areas after extensive training.

6.2.3 The integration of acquisition and retrieval processes in execution

Though only the acquisition processes in the early and late stages are discussed in the above sections, the acquisition and retrieval processes in the long-term memory cannot be isolated from each other in the execution process, as discussed in Section 3.1. In the above discussions, an emphasis is made on the comparison of the acquisition processes between the early and late stages. It should, then, be asked what would happen in the 'middle' stage between the early and late stages. Even though the acquisition processes in the early and late stages are discussed separately in the above discussions, the learning processes of the sensory-motor mapping, the sensory-sequence memory, and the motor-sequence memory concurrently occur at different learning speeds. The choice of an action to a state may happen to be different among these learning processes. In other words, different contexts should be distinguished from a computational viewpoint. It is, then, required to integrate these learning processes, as discussed in the beginning of Section 6.2.

Presumably, the sensory-sequence memory should be acquired faster than the motor-sequence memory as discussed in Section 5.3 as well as in the beginning of Section 6.2. It is postulated in this study that the sensory-sequence prediction controls the integration of decisions for a motor output among the sensory-motor mapping, the sensory-sequence memory, and the motor sequence memory. Sensory-sequence memory information can be propagated to the SMA through the pre-SMA, while the pre-SMA can pass its own output to the SMA as well.

6.3 A general framework on the acquisition and retrieval processes in execution

Computational elements in the basal ganglia-thalamocortical loops for sequential decision making are hypothesized in the previous section. Based on this proposal, the aim of this section is to provide a general framework describing the functions of the basal ganglia-thalamocortical loops for the acquisition and retrieval processes in execution. The emphasis is placed on ongoing processes in the early and late stages of learning sequential decision making.

1. **Exploration of the correspondence of the sensory inputs with the motor outputs**: In general, exploration to find optimal actions in relation to states is needed particularly in the early stage of the learning in sequential decision making.

In the 2x5 task, through the process of reinforcement leaning, the monkeys explore all possible pairs of *current* state and *current* action over all states. In this learning process, reinforcement learning will lead to the possibly optimal pairs of state and action being chosen more often. This function mapping the sensory input to the motor output is called *sensory-motor mapping* in the the rest of the thesis for convenience (Figure 6.1, top). The learning of the sensory-motor mapping is hypothesized to occur in the pre-SMA and the anterior striatum (Figure 6.2). The reinforcement signals are provided by the dopamine (DA) neurons in the substantia nigra pars compacta (SNc).

2. **Sensory-dependent learning of the sequence**: The learning of the sequences first occur depending upon the sensory inputs rather than the motor outputs. In the 2x5 task, it is particularly so because the learning of the sequence is much easier in Cartesian representations than in kinematical/dynamical representations. The memory of the sensory-dependent sequences helps the monkey anticipate the coming sensory input. This function is called *sensory-sequence prediction* in the rest of the thesis for convenience (Figure 6.1, middle). The learning of sensory-sequence prediction is hypothesized to occur in the dorsolateral prefrontal circuit (the dorsolateral prefrontal area (DLPF) and the anterior striatum; Figure 6.2). The reinforcement signals are also provided by the DA neurons. We hypothesized that the memory of the sensory-sequence prediction has two functional roles in execution: one is to evaluate how the expected context, or the expected sequence of sets in hyperset, fits well with the actual sequence of sets. The other is to anticipate the correct target corresponding to the anticipated sensory inputs, possibly in cooperation with the oculomotor circuit. The latter functional role may facilitate the success of the sequence and, consequently, facilitates the learning of the motor-

Figure 6.1

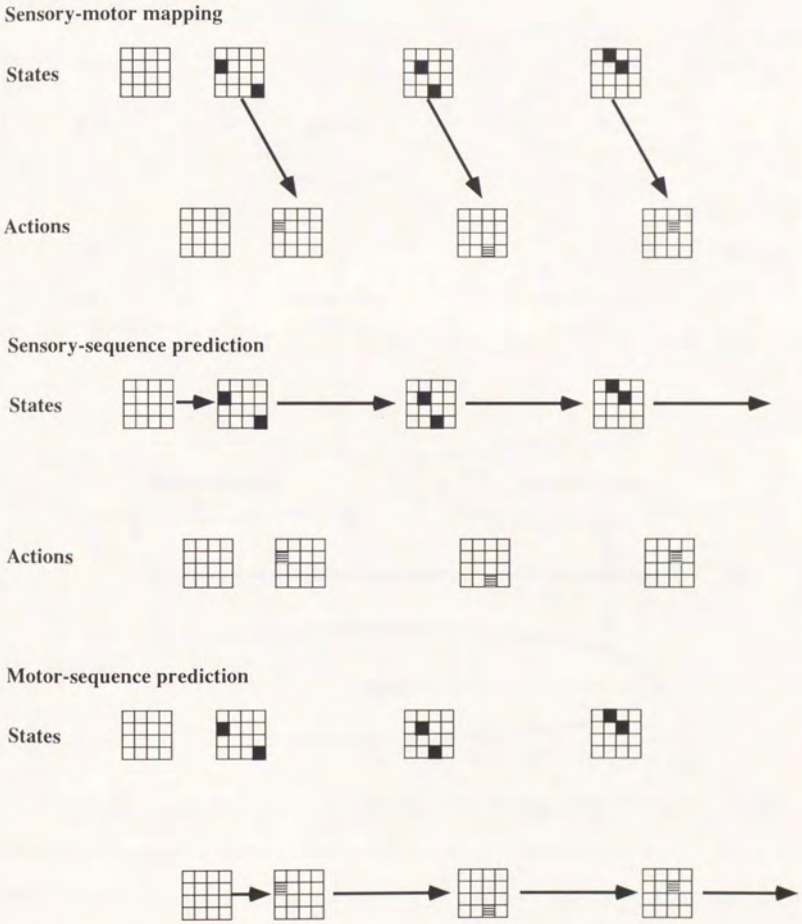


Figure 6.1: Sensory-motor mapping (top), sensory-sequence prediction (middle), motor-sequence prediction (bottom).

Figure 6.2

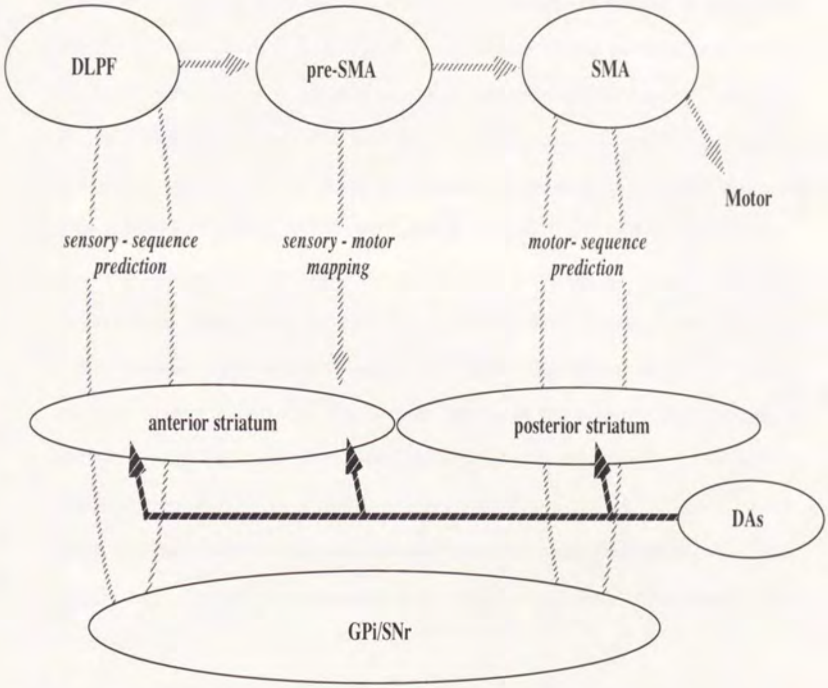


Figure 6.2: Correspondence of neural circuitry with the proposed functions. The dorsolateral prefrontal circuit corresponds with the sensory-sequence prediction. The pre-SMA with the anterior striatum corresponds with the sensory-motor mapping. The motor circuit corresponds with the motor-sequence prediction.

dependent sequences, which is discussed below.

3. **Motor-dependent learning of the sequence**: The memory of motor-dependent sequences helps the monkeys anticipate the correct action to the coming sensory input. This function is called *motor-sequence prediction* in the rest of the thesis for convenience (Figure 6.1, bottom). The learning of the motor-dependent sequence is postulated to occur more gradually than that of the sensory-dependent sequence (See the discussion in Section 6.2). The more the monkeys experience successes in the 2x5 task, the more the learning of the motor-dependent sequence is facilitated. Probably this slower learning process of the motor-dependent sequence is the reason that the performance time decreases more gradually than the number of trials in the 2x5 task. The learning of the sensory-motor mapping and the sensory-sequence prediction increase the number of successes of completed sets and trials. In this way, the learned results of the sensory-motor mapping and the sensory-sequence prediction can be gradually transferred to the motor-sequence prediction. This learning is hypothesized to occur in the motor circuit (SMA and the posterior striatum) with reinforcement signals from the DA neurons (Figure 6.2), probably in interaction with other motor-related cortical areas and the cerebellum.

4. **Integration of acquisition and retrieval processes in execution**: In the execution of the 2x5 task, even though the learning speeds are different, all learning processes simultaneously occur: the learning process of the sensory-motor mapping, that of the sensory-sequence prediction, and that of the motor-sequence prediction. Each of these learning processes can by itself lead to the determination of an action given a state. It is required, therefore, to integrate decisions made

by these three learning processes to produce a final decision of the action. As postulated in Section 6.2, the sensory-sequence prediction, hypothesized to occur in the dorsolateral prefrontal circuit, is responsible for this integration. The degree of correctness of the decisions made by each of these three learning processes may vary according to how often the sequence is experienced, or how well the given sequence is already learned by the sensory-sequence and motor-sequence predictions as well as the sensory-motor mapping.

6.4 A model of functions of the basal ganglia loops in the 2x5 task

Based on the general framework of the functions of the loops of the basal ganglia in the previous section, this section aims to provide a model implementing an algorithm, by use of which the computer simulation on the 2x5 task is employed in the next chapter.

First, several simplifications made in the construction of the model are stated and, then, a model implementing an algorithm based on the framework is explained.

6.4.1 Simplifications

Because the aim of the model in the following sections is intended to present a core of computational processes proposed in the preceding sections, there are several simplifications made in the construction of the model. The aim of this section is to state these simplifications. Figures in this section are presented in correspondence with Figures in the previous section in order to show what functions are implemented in the model among the functions hypothesized in the previous section.

Figure 6.3

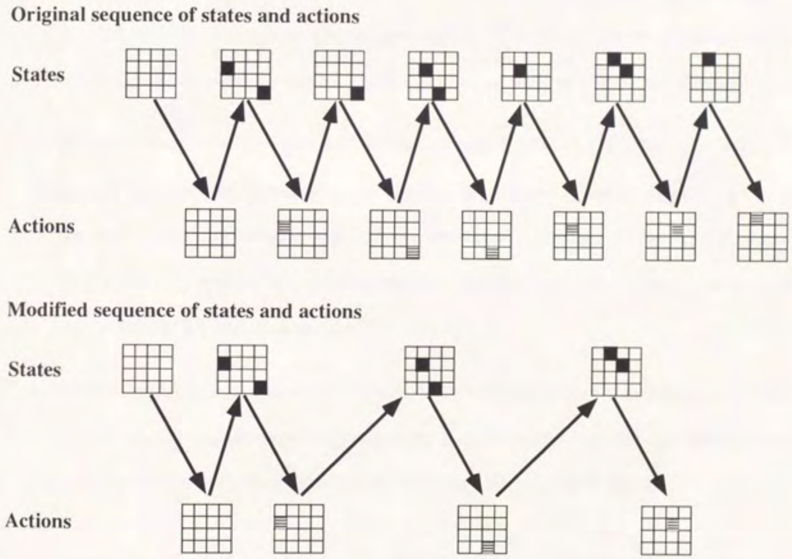


Figure 6.3: The sequence of states (sensory inputs) and actions (motor outputs) in the 2x5 task: the original sequence (above) and the modified sequence used in the simulation (below).

1. **Original Sequence of states and actions**: In the 2x5 task, or generally in the sequential decision making, the sequences of sensory inputs, which are also called states in particular in the framework of reinforcement learning (RL), and of motor outputs, which are also called actions in the framework of RL, are given as seen in Figure 6.3 (above). Because it is easy to push the second button in each set if the only one button is illuminating and because it is very rare for the monkeys in the 2x5 task to fail to push the second button in each set, the modified sequence of states and actions is used in the simulation, as seen in Figure 6.3 (below).

2. **Sensory-sequence prediction**: In the proposed scheme, though it is possible in general for the sensory-sequence prediction to rely on a whole sensory inputs in the past to predict the coming one, we assumed in the simulation as a minimal model that the sensory-sequence prediction uses the current and previous sensory inputs to predict the coming one (See Figure 6.4).

3. **Motor-sequence prediction**: Similar to the assumption of the sensory-sequence prediction, the motor-sequence prediction is assumed to use only the current and previous motor outputs to predict the coming one (Figure 6.4).

In the simulation of this study, the functional roles of the cerebellum is very simplified and the kinematical and dynamical aspect of motor control is neglected, in order to rather focus the coordination of the acquisition and retrieval processes in execution. Therefore, several experimental results in the 2x5 task measured by the performance time, one of measures used by Hikosaka and his colleagues, and other measures related to the motor control, are not examined in our simulation. This point will be discussed later in the next chapter in relation to the limitations

Figure 6.4

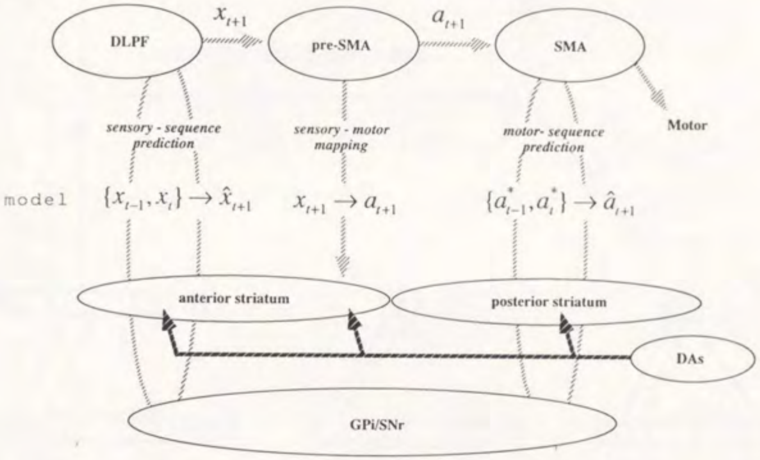


Figure 6.4: Correspondence between neural circuitry and the proposed functions, along with the functional relationship between a sensory input (x_t) and a motor output (a_t), which is implemented in the model. The dorsolateral prefrontal circuit corresponds with the sensory-sequence prediction. The pre-SMA with the anterior striatum corresponds with the sensory-motor mapping. The motor circuit corresponds with the motor-sequence prediction. Note that the notations in this figure are different from the notations that are used later in this chapter to explain an implemented algorithm in order to keep the simplicity in this figure.

Figure 6.5

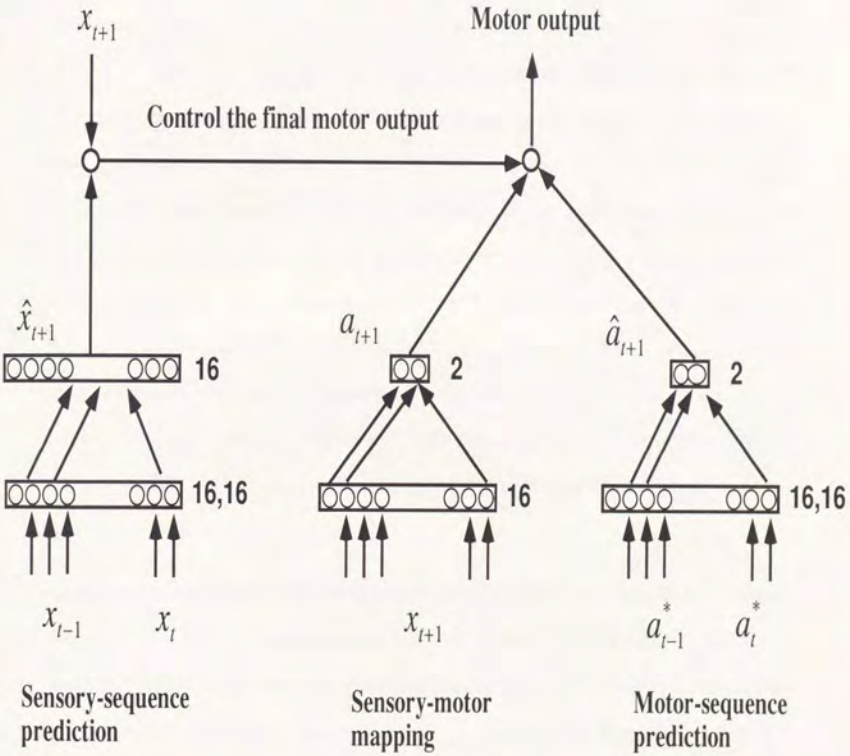


Figure 6.5: Functional relationship among the sensory-motor mapping, the sensory-sequence prediction, and the motor-sequence prediction, which is implemented in the model. x_t and a_t stands for a sensory input (state) and a motor output (action) at a time, t, respectively. Note that the notations in this figure are different from the notations that are used later in this chapter to explain an implemented algorithm in order to keep the simplicity in this figure.

and future works of this study.

We posited that both of the sensory-sequence prediction and motor-sequence prediction contribute to the anticipation of an action for a coming input. The anticipation by the sensory-sequence prediction will indirectly facilitate the learning of the motor-sequence prediction, if the anticipation by the sensory-sequence prediction increases the number of successes of the sequences. In terms of choosing an action, thus, the nature of the sensory-sequence and motor-sequence predictions is the same as *anticipating* an action in contrast to the sensory-motor mapping that *reacts* to a sensory input to give an action, and the facilitation of the learning of the motor-sequence prediction is only indirectly related to the sensory-sequence prediction. Based on these considerations, for the sake of simplicity, the anticipation of an action by the sensory-sequence prediction is only implicitly included in the algorithm given in the next section by letting the learning of the motor-sequence prediction advance relatively fast.

4. **Integration of acquisition and retrieval processes in execution**: Because the role of the sensory-sequence prediction for anticipating an action is implicitly included in the motor-sequence prediction in the algorithm, the issue of this integration is concerned only with the sensory-motor mapping and the motor-sequence prediction (Figure 6.5).

The following sections explain a model implementing an algorithm, based on the aforementioned general framework, which is employed in the simulation of the 2x5 task in the next chapter. Parameter settings and initializations, used in the simulation, are explained in the next chapter, Section 7.2.

In the following sections, each hyperset is denoted by h_i. In the 2x5 task, there are five sets in a hyperset. A set in a hyperset, h_i, is denoted by $x_t^{h_i}$. In a set, there are two illuminating buttons in the 4x4 matrix display. Hence, there are $4 \times 4 = 16$ buttons in the display. For the sake of simplicity, we regard a set, $x_t^{h_i}$, as a column vector of 16 dimensions, each element of which corresponds with the numbered button, and then each set can be represented as a vector of 16 dimensions with two 1s, corresponding to the place of two illuminating buttons, and other fourteen 0s. An action taken in each set is to push either of two illuminating buttons. When it is clear, we sometimes simplify the notation of a set, $x_t^{h_i}$, as x_t.

6.4.2 Sensory-motor mapping

The function of the sensory-motor mapping is modeled in the actor-critic scheme, which is discussed in Section 4.6. In the actor-critic scheme, the critic estimates the value function and the actor estimates the optimal policy as the exploration advances. In our scheme, the actor corresponds to the sensory-motor mapping, whereas the critic corresponds to the circuit of the striatum and SNc.

As seen in the discussion in Section 4.5, let us assume that the estimated value function by the critic, which is denoted by P, takes the linear function of the state, x_t, with the weight vector denoted by v. The function of the critic is, then, defined as

$$P(x_t) = \sum_{j=1}^{16} v_j (x_t)_j \tag{6.1}$$

The temporal difference(TD) error as seen below

$$\hat{r}_t = r_t + \gamma P(x_{t+1}) - P(x_t)$$

will be used to update the estimated value function of the critic by the method of the steepest descent as follows:

$$\begin{aligned}\Delta v_j &\propto \hat{r}_t \frac{\delta P}{\delta v_j} \\ &\propto \hat{r}_t (x_t)_j,\end{aligned}$$

where Δv_j stands for change of v_j by this update.

We assume the output of the sensory-motor mapping, that is, the actor, is given by a probability vector of 2 dimensions, given that the first element of the vector stands for the probability of pushing the smaller numbered button first among two illuminating buttons according to our numbered notation of 4x4 matrix display, and the second for the probability of pushing the larger numbered button first (Figure 6.5). Let us denote the output of the sensory-motor mapping by a^s. Since it is always true that the sum of the first and second elements of a^s is equal to 1, let us denote the first element of a^s by p_{a^s}. Then,

$$a^s = (p_{a^s}, 1 - p_{a^s}) \tag{6.2}$$

We assume that two actors compete with each other and each of them supports to push either the smaller or larger numbered button in our notation, given the information of a current set (See Section 4.6). Let us denote these two actors by $f_{a_1^s}$ and $f_{a_2^s}$ and outputs of them by q_{s1} and q_{s2}. Then,

$$q_{s1} = f_{a_1^s}(x) \tag{6.3}$$

$$q_{s2} = f_{a_2^s}(x), \tag{6.4}$$

where x is the current set. For the sake of simplicity, we defined $f_{a_n^s}$ $(n = 1, 2)$ as the inner product of the vector of the current set with the sixteen-dimensional weight vector, termed as $W_{a_n^s}$ $(n = 1, 2)$. Then,

$$q_{s1} = W_{a_1^s} \cdot x \tag{6.5}$$

$$q_{s2} = W_{a_2^s} \cdot x \tag{6.6}$$

To give the probability of pushing either button based on outputs of two actors, q_{s1} and q_{s2}, we introduce the sigmoid function, F, as follows:

$$F(z) = \frac{1}{1 + exp(-z)}, \tag{6.7}$$

where z is a real number.

Then, p_{a^s} is defined as follows:

$$p_{a^s} = F(q_{s1} - q_{s2}) \tag{6.8}$$

6.4.3 Sensory-sequence prediction

To predict the coming sensory input, the sensory-sequence prediction uses the information of the current set and the set before the current. Let us denote the function of the sensory-sequence prediction and the output of the sensory-sequence prediction by f_s and $\hat{x}$, respectively. Then, the sensory-sequence prediction is defined as:

$$\hat{x}_{t+1} = f_s(x_t, x_{t-1}) \tag{6.9}$$

f_s is a function that has the mapping: $16 \times 16 \rightarrow 16$.

Because there will be used a few bilinear forms in the proposed algorithm including the sensory-sequence prediction, as seen below, let us define a bilinear form with two vector inputs, p, q, for convenience, as follows:

$$G(p, q) = p^T G q, \tag{6.10}$$

where T of p^T represents the transpose.

Then we define the function of the sensory-sequence prediction, f_s, using the sigmoid function defined in the previous section, F, as follows:

$$\begin{aligned}(\hat{x}_{t+1})_j &= f_s(x_t, x_{t-1}) \\ &= F(G_{C_j}(x_t, x_{t-1})) \qquad (j = 1, 2, ..., 16) \qquad (6.11)\end{aligned}$$

where G_{C_j} is 16×16 matrix, taking bilinear with respect to x_t and x_{t-1}.

Learning of the sensory-sequence prediction

The sensory-sequence prediction anticipates the coming sensory input (the coming set) with information of the current set and the set before the current. The learning of the sensory-sequence prediction occurs with respect to the coming set as the target, when the monkeys did not fail to give a correct response to a current set, which actually means that a monkey could advance to the coming set. The reinforcement signal is assumed to be provided by dopamine neurons. The learning procedure follows a typical learning method of artificial neural network, that is, the method of the steepest descent with respect to the mean-squared error function [23].

$$\begin{aligned}\Delta(G_{C_j})_{kl} &\propto -\frac{\partial \frac{1}{2}((x_{t+1})_j - (\hat{x}_{t+1})_j)^2}{\partial (G_{C_j})_{kl}} \\ &\propto ((x_{t+1})_j - (\hat{x}_{t+1})_j)\frac{\partial(\hat{x}_{t+1})_j}{\partial(G_{C_j})_{kl}} \\ &\propto ((x_{t+1})_j - (\hat{x}_{t+1})_j)\frac{\partial F}{\partial G_{C_j}}\frac{\partial G_{C_j}}{\partial(G_{C_j})_{kl}} \\ &\propto ((x_{t+1})_j - (\hat{x}_{t+1})_j)(\hat{x}_{t+1})_j(1-(\hat{x}_{t+1})_j)(x_t)_k(x_{t-1})_l, \qquad (6.12)\end{aligned}$$

where indices k, l denote kth and lth element of the vectors and the kl element of the matrix.

6.4.4 Motor-sequence prediction

In the proposed algorithm, the motor-sequence prediction anticipates an action to the coming sensory inputs(the next set) based on the actions in the previous set and the set before the previous. There is, however, a technical problem to construct the algorithm in this model. The motor output in the algorithm is assumed to be 2 dimensional probability vector that provides the probabilities of pushing either the smaller numbered or the larger numbered button. Because the original space of the motor output is 16 dimension (there are 16 buttons that is potentially to be pushed), the current format of the motor output is severely degenerated. This makes the learning of action sequences much more difficult if the aforementioned 2 dimensional format of motor output is also used as the input format of actions. This difficulty is caused technically so that it is not worth involving it in the simulation. We used, therefore, the 16 dimensional vector of actions as inputs to the motor-sequence prediction, instead of the 2 dimensional vector which was used for the action as outputs. What are actually actions for the monkeys in each set in the 2x5 task? Actions in each set are to push the first button and then to push the second button in a correct order. Thus, we introduce a^* to represent actions in each set by 16 dimension, given that the taken actions in each set, the first button to push and then the second button to push. In other words, an action in a set, a^*, is a 16 dimensional vector with two 1s at the corresponding numbered elements for the first and second buttons and other 0s.

The output of the motor-sequence prediction takes a form similar to the sensory-motor mapping. Let us denote the output of the motor-sequence prediction by a^m, and the first element of the motor-sequence prediction by p_{a^m}. Then,

$$a^m = (p_{a^m}, 1 - p_{a^m}), \tag{6.13}$$

where p_{a^m} satisfies $0 \leq p_{a^m} \leq 1$.

Similar to the sensory-motor mapping, the value of p_{a^m} is given by use of the sigmoid function, F, as follows:

$$p_{a^m} = F(q_{c1} - q_{c2}), \tag{6.14}$$

where q_{m1} and q_{m2} is defined as follows:

$$q_{mn} = G_{A_n}(a_t^*, a_{t-1}^*) \qquad (n = 1, 2) \tag{6.15}$$

where G_{A_n} is 16×16 matrix.

Learning of the motor-sequence prediction

In brief, the definition of the motor-sequence prediction is

$$a^m = (p_{a^m}, 1 - p_{a^m})$$

where

$$p_{a^m} = F(G_{A_1}(a_t^*, a_{t-1}^*) - G_{A_2}(a_t^*, a_{t-1}^*)),$$

When a monkey did not fail to give a correct response to a set, which actually means that a monkey could proceed to the next set, learning of the motor-sequence prediction occurs with respect to a finally taken action. Though the final action probability vector is given by a_n^f, where $n = 1$ or $n = 2$, after actually taking an action, the probability to take an action should be set as 1. The method of the steepest descent, then, can be employed, similar to the sensory-sequence prediction, as follows:

$$\begin{aligned} \Delta(G_{A_n})_{kl} &\propto -\frac{\partial \frac{1}{2}((1 - a_n^m)^2}{\partial (G_{A_n})_{kl}} \\ &\propto (1 - a_n^m) \frac{\partial a_n^m}{\partial (G_{A_n})_{kl}} \end{aligned} \tag{6.16}$$

For $n = 1$,

$$\begin{aligned}
\Delta(G_{A_1})_{kl} &\propto (1 - a_1^m)\frac{\partial a_1^m}{\partial(G_{A_1})_{kl}} \\
&\propto (1 - a_1^m)\frac{\partial p_{a^m}}{\partial(G_{A_1})_{kl}} \\
&\propto (1 - a_1^m)\frac{\partial F}{\partial G_{A_1}}\frac{\partial G_{A_1}}{\partial(G_{A_1})_{kl}} \\
&\propto (1 - a_1^m)p_{a^m}(1 - p_{a^m})(a_t^*)_k(a_{t-1}^*)_l \\
&\propto (1 - p_{a^m})p_{a^m}(1 - p_{a^m})(a_t^*)_k(a_{t-1}^*)_l \qquad (6.17)
\end{aligned}$$

For $n = 2$,

$$\begin{aligned}
\Delta(G_{A_2})_{kl} &\propto (1 - a_2^m)\frac{\partial a_2^m}{\partial(G_{A_2})_{kl}} \\
&\propto (1 - a_2^m)\frac{\partial(1 - p_{a^m})}{\partial(G_{A_2})_{kl}} \\
&\propto (1 - a_2^m)\frac{\partial - F}{\partial G_{A_2}}\frac{\partial G_{A_2}}{\partial(G_{A_2})_{kl}} \\
&\propto (1 - a_2^m)p_{a^m}(1 - p_{a^m})(a_t^*)_k(a_{t-1}^*)_l \\
&\propto p_{a^m}p_{a^m}(1 - p_{a^m})(a_t^*)_k(a_{t-1}^*)_l \qquad (6.18)
\end{aligned}$$

Therefore, we can summarize for both $n = 1, 2$,

$$\Delta(G_{A_n})_{kl} \propto (1 - a_n^m)p_{a^m}(1 - p_{a^m})(a_t^*)_k(a_{t-1}^*)_l \qquad (6.19)$$

6.4.5 Integration of sensory-motor mapping and motor-sequence prediction

As discussed in Section 6.3, the integration of the sensory-motor mapping and motor-sequence prediction is needed. The sensory-sequence prediction is postulated to control this integration. The degree to which the anticipated coming sensory inputs by the sensory-sequence prediction match with the actual coming sensory inputs is considered as the control of the integration. This corresponds to the situation in the experiment of the 2x5 task such that when the monkeys can anticipate well the coming sensory inputs, the output of the motor-sequence prediction is more reliable and, in contrast, that when the coming sensory input is different from the anticipated one, the output of the motor-sequence prediction is less reliable. For this reason, the concept of the *biasing* is introduced. When the value of biasing is high, the monkeys have predicted well the coming sensory inputs and vise versa. Let us denote the value of biasing by b_{x_t} given a set, x_t and denote the function of biasing by f_b. Then,

$$b_{x_t} = f_b(x_t, \hat{x}_t) \tag{6.20}$$

where $\hat{x}_t$ is the output of the sensory-sequence prediction.

We denote the final output by the integration of the sensory-motor mapping and motor-sequence prediction and the first element of the final output by a^f and p_{a^f} respectively. Then, we define a^f and p_{a^f} as follows:

$$a^f = (p_{a^f}, 1 - p_{a^f}) \tag{6.21}$$

$$p_{a^f} = \begin{cases} p_{a^m} & \text{with the probability } b_x \\ p_{a^s} & \text{with the probability } 1 - b_x \end{cases} \tag{6.22}$$

To make the biasing, b_{x_t}, concrete, we define the function of the biasing, f_b, as follows:

$$\begin{aligned} b_{x_t} &= f_b(x_t, \hat{x}_t) && (6.23) \\ &= \frac{x_t \cdot \hat{x}_t}{x_t \cdot x_t}, && (6.24) \end{aligned}$$

In the above equation, by the definition of the sensory-sequence prediction discussed before, the following is true:

$$0 \leq (\hat{x}_t)_j \leq 1 \qquad (j = 1, 2, ..., 16) \tag{6.25}$$

Therefore, the following also holds:

$$0 \leq b_{x_t} \leq 1 \tag{6.26}$$

6.4.6 Special cases for the first and second sets in the model

In the simulation of the 2x5 task, the definitions of the sensory-sequence prediction and motor-sequence prediction are introduced to reflect the fact that the monkeys gradually develop the anticipation of the coming sensory inputs and its corresponding motor outputs over a long period and can do so well particularly for extensively experienced hypersets (learned hypersets), as shown in Miyashita et. al. [45]. We first assumed that an anticipation of a set, x_{t+1} is based on the information of any of sets and actions a monkey previously experienced, $\{x_l\}_{l=1}^{t}$ and $\{a_l\}_{l=1}^{t}$. The simplest one is, then, chosen such that a monkey uses the information of sensory inputs and motor outputs in the previous set and the set before the previous, as explained in the preceding sections. In this definition, there is a problem for the first and second sets so that we need a special treatment for these cases. First, the case of the second set is explained, and then, the case of the first set is explained.

Special case for the second set

If we take our definition of the sensory-sequence and motor-sequence predictions literally, the inputs to the sensory-sequence and motor-sequence predictions should be the sensory inputs and motor outputs in the first set and the ones in the set before the first. The latter does not exist. To avoid this technically, it is assumed that both of two inputs to the sensory-sequence and motor-sequence predictions are the ones in the first set, that is, (x_1, x_1) and (a_1^*, a_1^*), respectively.

Special case for the first set

The sensory-sequence and motor-sequence predictions, as discussed so far, anticipate the coming sensory inputs and its motor outputs based on the sensory and motor information in the current set and the set before the current. The information *within a hyperset*, thus, is used for this anticipation. In other words, the monkeys develop the contextual decision with respect to each hyperset, using information *within a hyperset.* In the currently chosen algorithm, the monkeys cannot use the sensory-sequence and motor-sequence predictions to anticipate the first set and its corresponding action in a hyperset. It is, however, very likely that there may exist a larger context by which the monkeys can tell that they are going to engage a different hyperset between blocks in the experiment. In this case, experiences of hypersets, in particular of learned hypersets, are likely to influence its decision. Thus, it is plausible to assume to some degree that there may exist the context even for the first set.

There can be no decision made by the sensory-sequence prediction/motor-sequence prediction in the first set. Furthermore, we assume in our simulation that,*each time between blocks, the weight matrices of the sensory-motor mapping and the critic is initialized.* Therefore, there is no learning effect on the selection of actions by the sensory-motor mapping in any set, including the first set, before it learns. The contextual effect for the first set that should be taken into account, therefore, is that the monkeys should develop preference to choose the actions in the first set, based on the past experience of the 2x5 task, particularly that of learned hypersets,.

Recall that the output of the sensory-motor mapping is defined as

$$\begin{aligned} a^s &= (p_{a^s}, 1 - p_{a^s}) \\ p_{a^s} &= F(W_{a_1^s} \cdot x - W_{a_2^s} \cdot x), \end{aligned}$$

where F is the sigmoid function.

We assume that in the first set, there is the input of 16 dimensional vector, which is the first set, x_1, denoted by x_{start}, to tell a monkey to start the first set and that this information is taken into account in the above formula with the 16 dimensional weight vector, denoted by W_{start}, as follows:

$$p_{a^s} = F(W_{a_1^s} \cdot x - W_{a_2^s} \cdot x + W_{start} \cdot x_{start}) \tag{6.27}$$

When the choice of the action was correct, that is, when a monkey succeeds to proceed to the second set, the learning of W_{start} occurs, similar to that of the motor-sequence prediction, as follows:

$$\Delta(W_{start})_j \propto \begin{cases} p_{a^s}(1 - p_{a^s})x_j & \text{when the action, } a_1\text{, is taken} \\ -p_{a^s}(1 - p_{a^s})x_j & \text{when the action, } a_2\text{, is taken} \end{cases} \tag{6.28}$$

In this way, W_{start} can reflect the contextual effect on the action selection in the first set, which is based on overall experiences in the 2x5 task.

Chapter 7

SIMULATION OF THE 2X5 TASK

7.1 Introduction

In this chapter, by use of the model proposed in the previous chapter, the simulation of the 2x5 task is performed. The results of the simulation is compared closely with the experimental data of Hikosaka laboratory, followed by the discussion.

7.2 Simulation methods

7.2.1 Parameter settings

Initialization of weight parameters

Generally speaking, the initial weights of functions make an influence on the learning progress in reinforcement learning, though its influence is linked with the magnitude of the learning parameters. Sometimes, it happens based on the combination of the initial weights and learning parameters that the learning becomes severely slow. The strategy taken for parameter tuning in this study is, therefore, that the initial weights are, first, set prior to deciding learning parameters, according to the consideration the monkeys in the 2x5 task as explained below. Then, learning parameters are optimized with those initial weights, as seen in the next section.

All elements of the initial weights of the sensory-motor mapping, the sensory-sequence prediction, and the motor-sequence predictions are set as 0. This is because we assumed that the monkeys do not have any prejudice to push any button before experiencing any of hypersets. All elements of the weight of the value function is set as 0.5. The monkeys exhibited the long-term and sequence-nonselective learning after experiencing the 2x5 task to some degree. We assumed that this was because the monkeys have got some expectation to receive the rewards by situating himself/herself in the 2x5 task. So 0.5 is chosen rather than 0.0.

For the value function, rewards should be determined for each transition of states. Specifically speaking, in the 2x5 task, rewards should be determined in cases such that the monkeys succeed to proceed from the 1st to the 2nd (denoted by r_1), from the 2nd to the 3rd (r_2), from the 3rd to the 4th (r_3), from the 4th to the 5th (r_4) and succeed with the 5th set (r_5), and, in addition, when the monkeys fail to proceed at each set (denoted by r_f). Let us denote the set of rewards, $\{r_1, r_2, r_3, r_4, r_5, r_f\}$ by $\mathbf{r}$.

In the experiment of the 2x5 task, the monkeys have received a small amount of rewards between sets, though the amount would vary in each transition and, the amount is typically the highest when the monkeys succeed a trial. The reason that a small amount of rewards are given in each transition between sets is primarily because of the training. Without it, the monkeys would simply lose the interest in doing the task particularly when they do not have much experience with the task.

We tested several $\mathbf{r}$ such as $\{0.2, 0.2, 0.2, 0.2, 0.2, -0.2\}$, $\{0.2, 0.2, 0.2, 0.2, 0.2, 0.0\}$, $\{0.0, 0.0, 0.0, 0.0, 1.0, -0.2\}$, and $\{0.0, 0.0, 0.0, 0.0, 1.0, 0.0\}$. In a certain range of the learning parameters of the critic and the sensory-motor mapping (the actor), the learning converged almost at the same rate. With the learning parameters of those in that range, $\mathbf{r} = \{0.2, 0.2, 0.2, 0.2, 0.2, -0.2\}$ is used in the simulation reported in this study.

γ should be decided to construct the temporal difference error. It is observed in the experiment that the monkeys receive a higher (or at least the same) amount of the reward as the monkeys advance in each trial, and that, as in the experimental situation, the monkeys are interested to some degree in completing a trial after a certain experience of task. The latter fact is because the monkeys keep doing a task for a while even without the rewards, though they stop doing the task after a while if the rewards are still absent. Therefore, it is better to keep $\gamma > 0$, and we chose $\gamma = 0.5$.

Learning parameters

In the proposed algorithm, there are a multiple of learning processes such as for the value function, the sensory-motor mapping, the sensory-sequence prediction, and the motor-sequence prediction. There, thus, are a multiple of learning parameters for each of these functions. Changing the magnitude of learning parameters makes an effect on the learning processes.

First, without the sensory-sequence prediction and the motor-sequence prediction, we tested many different sets of the learning parameters of the value function and the sensory-motor mapping. By comparing the performance of these parameters, the learning parameters of the value function and the sensory-motor mapping were then determined as nearly optimal, though it is impossible to determine the precise optimal parameters in general (See Appendix C for the actual numbers). Secondly, with these learning parameters of the value function and the sensory-motor mapping, many different sets of the learning parameters of the sensory-sequence prediction and the motor-sequence prediction were tested. Again, these learning parameters were determined as nearly optimal (See Appendix C for the actual numbers).

7.2.2 Training of learned hypersets

In the actual experiment of the 2x5 task, from 14 up to 20 hypersets are used to train the monkeys each day. Most of hypersets are presented to the monkeys only once, called *new hypersets*. Some of hypersets are extensively experienced almost every day in a certain period and, then, the monkeys become very good at performing these hypersets (See Section 5.2.1 and Section 5.3). These hypersets are called *learned hypersets* after such extensive experiences along with the improvement of the performance [25]. The hypersets that are used for such extensive experiences and become the learned hypersets are chosen in advance in the experimental procedure of Hikosaka et al. [25].

In the experiment of the 2x5 task, among 14 up to 20 hypersets per a day, more than half of them, at least the half of them, are learned hypersets. It is, therefore, considered that the acquired ability of performing the learned hypersets is not interfered much by the learning of new hypersets, though the interference should exist at least. In addition, the weights of the sensory-motor mapping and the critic are initialized at the beginning of each block in the simulation. For the sake of simplicity, therefore, the training procedure in the simulation is simplified as explained below.

In the following sections, the pre-chosen hypersets that are to become learned hypersets are called learned hypersets, even before the training, in the explanation of the simulation for convenience. 10 learned hypersets are generated by the procedure that is explained in the next section. The model experienced all of 10 learned hypersets in a randomized order each day. The training has been done for 30 days. After the training, the weights of the sensory-sequence prediction and the motor-sequence prediction is preserved to test several conditions.

7.2.3 Generation of hypersets

10 hypersets that are used for the training of the proposed model, called *learned hypersets*, are generated randomly with several constraints described below:

1. It is assumed that there is no same set within a hyperset, following the experimental setting of the 2x5 task (Miyachi, personal communication).
2. Because we use the minimal model of the sensory-sequence and motor-sequence prediction in that these predictions only use the information of the current set and the set before the current set to predict the coming sensory input and its corresponding action respectively, it is assumed that among all hypersets, there are no two same sets in the same order.
3. Because the sensory-sequence and motor-sequence prediction can not be applied fully for the first set, as discussed in Section 6.4.6, it is assumed that there appears no same set as the first set among both of learned and new hypersets.

Following the same constraints, a few tens of other hypersets are also generated, called *new hypersets*, which are used to examine the performance in several conditions with the weights of the sensory-sequence and motor-sequence predictions obtained after the training.

All of the *learned hypersets* are shown in Figure 7.1.

Figure 7.1

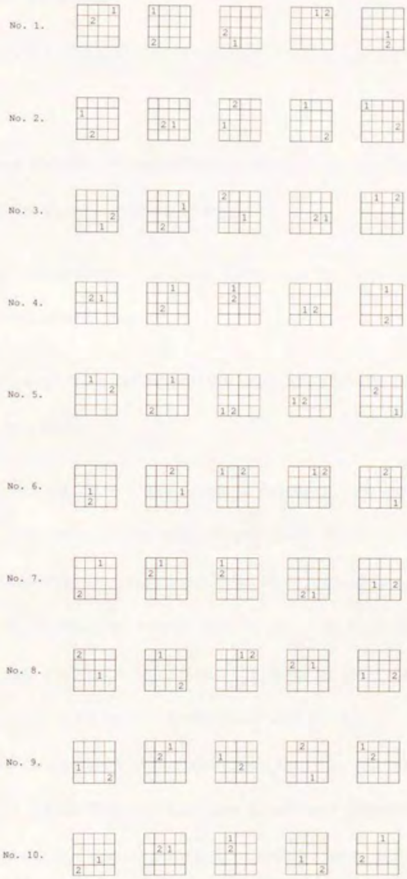


Figure 7.1: Hypersets are used in the simulation, all of which became the learned ones after 30 days training.

7.3 Results

7.3.1 Learning process

As discussed in Section 5.3, Hikosaka et al. [25] reported three different levels of learning as follows:

1. **short-term and sequence-selective**: indicated by improved performance for a particular hyperset *during a block of trials.*
2. **long-term and sequence-selective**: indicated by improved performance for a particular hyperset *across days.*
3. **long-term and sequence-unselective**: indicated by the improvement of performance for new hypersets.

In our simulation, among these three levels of learning, two levels, **short-term and sequence-selective** and **long-term and sequence-selective levels**, are examined. The learning parameters were optimized prior to the training in our simulation and, in addition, the cerebellar components were also omitted in the model used in the simulation. It is, therefore, irrelevant to discuss the level of **long-term and sequence-unselective learning** using the present model of this study.

Hikosaka et al. [25] have shown experimentally that the monkeys make more errors in the early stage of a block than in the late stage and concluded that it indicates learning at the short-term and sequence-selective level (Figure 7.2). The same tendency is observed in the performance of the model. An example of the performance of the model for several (simulated) days is shown in Figure 7.3 for a hyperset (No.9 in Figure 7.1; Refer to Appendix D to see the performance on other hypersets.) for the 1st day (top), the 3rd day (middle), and the 30th day (bottom). It is clear that on each day, the

Figure 7.2

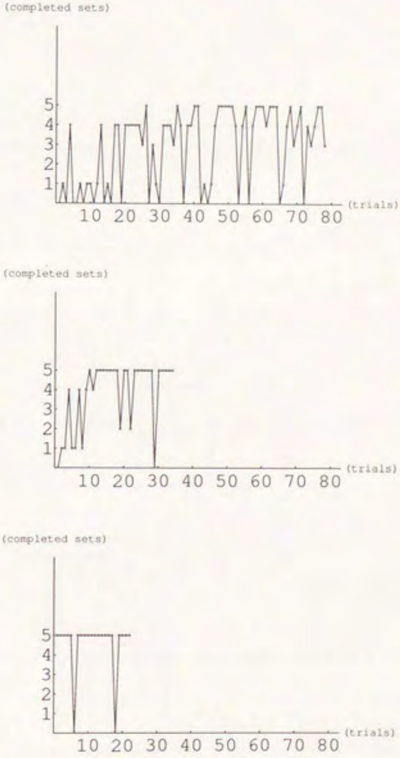


Figure 7.2: Experimental result of learning a hyperset across trials. The change in the number of completed sets (ordinate) across trials (abscissa) is compared among the 1st day (top), the 3rd day (middle) and the 30th day (bottom). Taken from Hikosaka et al. [25]

Figure 7.3

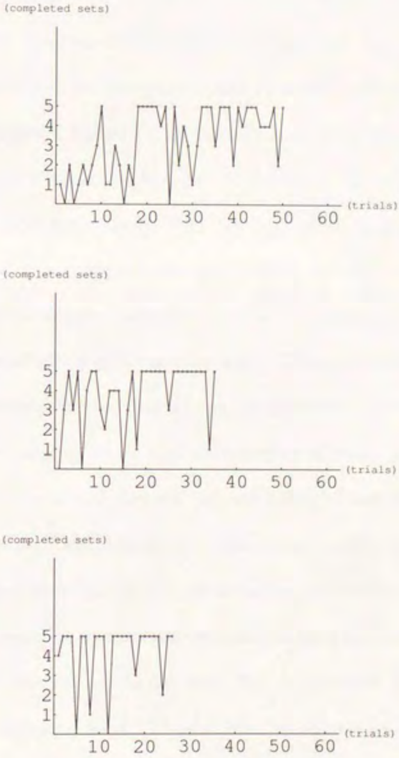


Figure 7.3: Example of learning a hyperset (No.9) across trials by the model. The change in the number of completed sets (ordinate) across trials (abscissa) is compared among the 1st day (top), the 3rd day (middle) and the 30th day (bottom).

errors are made more in the early stage than in the later stage. This indicates that the model exhibited the short-term and sequence-selective learning. Another example of the simulation is given in Figure 7.4 (No.2 hyperset; Refer to Appendix D to see the performance on other hypersets). In Figure 7.4, the cumulative number of successful trials is plotted against the sequential number of trials for a learned hyperset (No.7). Following Hikosaka et al. [25], the number of trials criterion is set as 10 successful trials, indicated by the broken line in Figure 7.4. It can be observed in Figure 7.4 that the model took a greater number of trials in the first half 10 successful trials, which is the number of trials (abscissa) at the crossing point with the horizontal broken line in Figure 7.4, than in the second half 10 successful trials, particularly for the 1st and 3rd days. The learning at short-term and sequence-selective level is particularly evident for the first 5 days, confirmed as statistically significant ($p < 0.05$). This result coincides well with the experimental result of Hikosaka et al [25], as seen in Figure 7.5, and indicates short-term and sequence-selective learning. Note that the number of trials to reach criterion (10 successful trials) for a block of each day will be used below to assess the learning process across days, as in the experimental study by Hikosaka et al. [25].

In Figure 7.4, an improvement of the performance across days, that is, long-term and sequence-selective learning, can be also found by a leftward shift and steepening of the learning curve from the 1st day to the 30th day. In addition, Figure 7.6 shows the number of trials to complete a block (20 successful trials) across days and Figure 7.7 shows the averaged number across days. It is obvious that the learning is more rapid in a first few days and, then, gradually asymptotes in both figures. This also indicates long-term and sequence-selective learning. Figure 7.8 shows the actual experimental data on the time course of learning of the monkeys, taken from Hikosaka et al. [25] with a slight modification. The performance of the model captures the characteristic of the

Figure 7.4

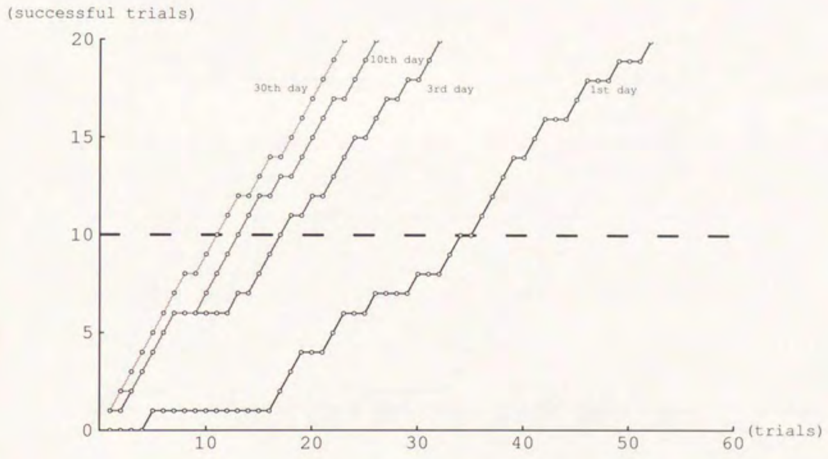


Figure 7.4: Example of learning a hyperset (No.2) across days by the model. The number of successful trials (ordinate) is plotted against the total number of trials (abscissa) for the 1st day, the 3rd day, the 10th day, and the 30th day from right to left.

Figure 7.5

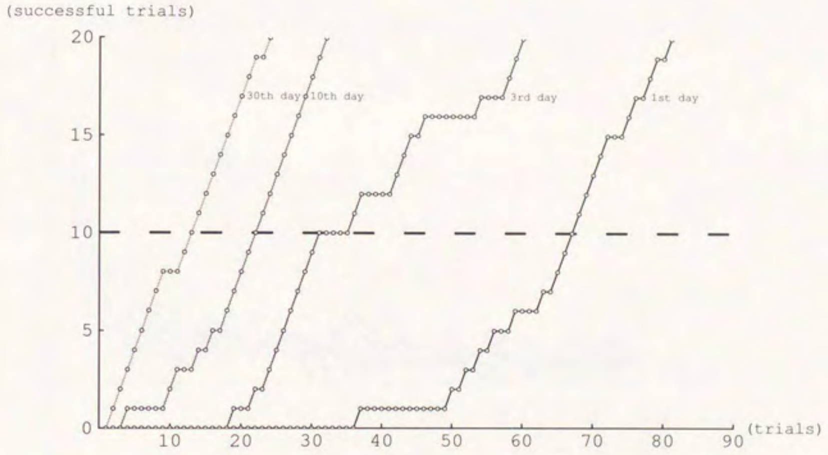


Figure 7.5: Experimental data on learning a hyperset across days. The number of successful trials (ordinate) is plotted against the total number of trials (abscissa) for the 1st day, the 3rd day, the 10th day, and the 30th day from right to left. Taken with a slight modification from Hikosaka et al. [25].

Figure 7.6

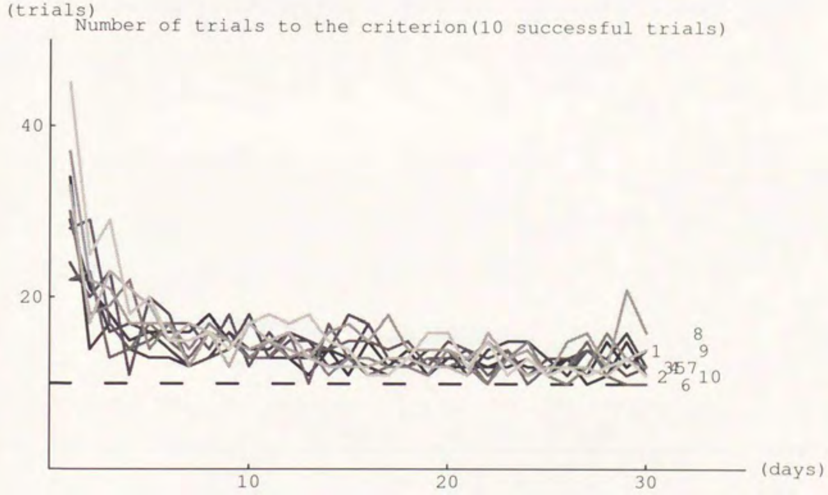


Figure 7.6: Learning of hypersets used in the training (learned hypersets) across 30 days. The number of trials to criterion (ordinate), that is, 10 successful trials, is plloted across days by the model (abscissa) for all 10 hypersets.

Figure 7.7

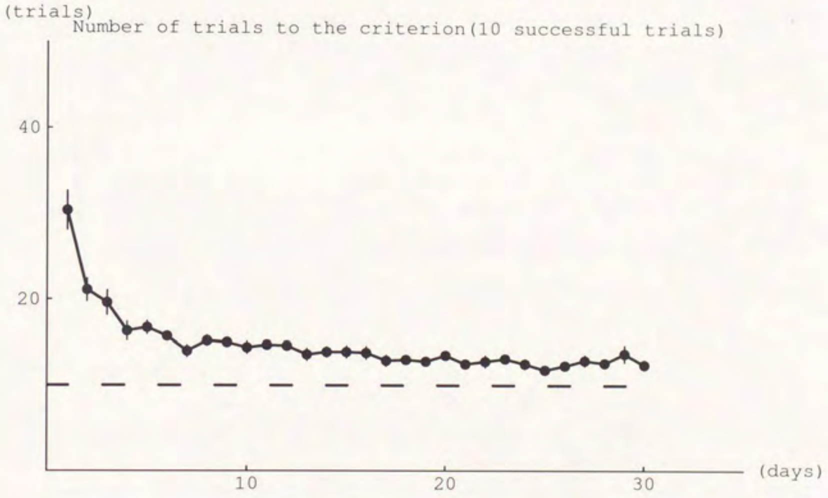


Figure 7.7: Averaged learning of hypersets used in the training (learned hypersets) across blocks over 30 days. The mean number of trials to criterion (ordinate) for all 10 learned hypersets is plotted across days by the model (abscissa). The bar at each data point represents the standard error (SE).

Figure 7.8

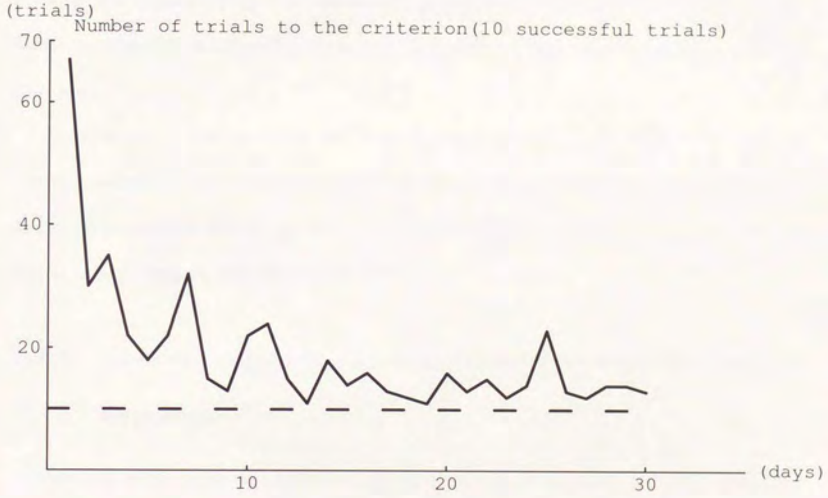


Figure 7.8: Experimental data on learning a hyperset across 30 days. The number of trials to criterion (10 successful trials) (ordinate) is plotted across days (abscissa). Taken with a slight modification from Hikosaka et al. [25]

experimental data that the number of trials rapidly decrease in a first few days and, then, gradually asymptotes. There is a difference in the actual number of trials between the experimental data and that of the simulation. This difference by itself, however, is not the issue, because the learning parameters of the model are optimized prior to the simulation and, therefore, the model should be examined for the qualitative behavior. In addition, by the third level of the learning, that is, long-term and sequence-unselective learning, the number of trials in the actual experiments tends to decrease, in particular for the first few blocks of each hyperset, as the monkeys have more experiences with the 2x5 task.

In summary, it can be concluded from the above results that the model exhibited learning not only at the short-term and the sequence-selective level but also at long-term and sequence-selective level, qualitatively similar to the learning of the monkeys in the actual experiments by Hikosaka et al. [25].

7.3.2 Context dependency of memory retrieval for learned hypersets

Hikosaka et al. [27] have compared the performance on the learned hypersets with that of *reversed hypersets*, that is, the hyperset of which all sets are the same as the learned hypersets but the sequence of the sets is reversed, as discussed in Section 5.3 (also see Figure 5.2). The results measured by the number of trials to criterion (10 successful trials) were that (1) the performance of the reversed hypersets was significantly worse than that of the learned hypersets and (2) the performance of the reversed hypersets was not significantly different from that of the new hypersets, whereas the performance of the learned hypersets was significantly better than that of both of the reversed and new hypersets.

Figure 7.9

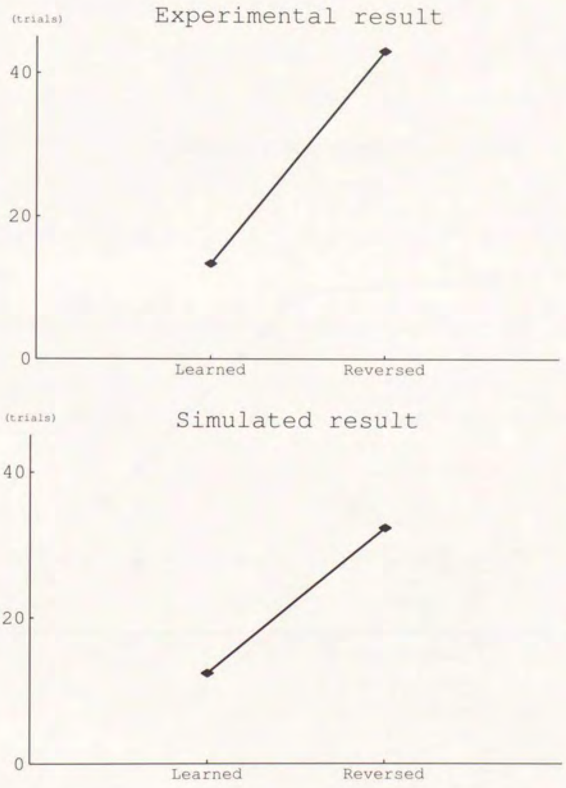


Figure 7.9: Averaged number of trails to criterion (10 successful trials) for the learned hypersets and the reversed hypersets: the results from the monkeys taken from Hikosaka et al. [27] (above) and the results from the model (below).

Figure 7.10

Figure 7.10: Averaged number of trails to criterion (10 successful trials) of the model for the learned hypersets (left), the reversed hypersets (middle), and the new hypersets (right). The bar at each data point represents the standard error (SE).

We tested the reversed hypersets for all 10 learned hypersets with the weights of the sensory-sequence prediction and motor-sequence prediction that were obtained after the training. The results showed the same characteristics as the actual monkeys in Hikosaka et al. [27] (Figure 7.9)[1]

By the measure of the number of trials to criterion (10 successful trials), the performance of the model for the learned hypersets (the mean: 12.4) was statistically significantly better than that for the reversed hypersets (the mean: 32.4, $p < 0.001$) and that for the new hypersets (the mean: 33.8, $p < 0.0005$). The performance for the reversed hypersets was not significantly different from that for the new hypersets ($p > 0.3$) (Figure 7.10).

The performance of the model, thus, coincides with that of the monkeys in conditions of the learned, reversed, and new hypersets. This result indicates that for the learned hypersets, the model chooses a motor output for each set, not depending upon the current sensory input (the current set) but rather depending upon the sensory inputs before the current sensory input. In other words, the model anticipates an action for the coming input and, in this way, the choice of actions by the model for the learned hypersets exhibited the context dependency similar to the monkeys.

7.3.3 Blockade of the sensory-motor mapping

Miyachi et al. [44] have shown that when the anterior striatum is blocked by muscimol (a GABA agonist) injection, the performance of the monkeys in the 2x5 task is significantly worse for new hypersets than that of the control condition (See Figure 7.11; also refer

---

[1] Because Hikosaka et al. [27] provide the actual number of the experimental data only in the conditions of the learned and reversed hypersets, not in that of the new hypersets, Figure 7.9 shows the comparison of only these conditions between experimental and simulated results.

to Section 5.4.1). In contrast, the performance for learned hypersets is not significantly different between the control and blockade conditions. Miyashita et al. [46, 47] have shown that when the pre-SMA is blocked by muscimol injection, the performance is significantly worse for new hypersets than that of the control condition, whereas it is not significantly different for the learned hypersets between the control condition and the blockade condition (See Figure 7.12. Also refer to Section 5.4.2.)

As postulated in Section 6.3, the pre-SMA functions implements the sensory-motor mapping in interaction with the anterior striatum in our proposed scheme. The anterior striatum is assumed to contribute to sensory-sequence prediction as part of the dorsolateral prefrontal circuit as well. It is, however, questionable to what degree the striatum is involved in storing information of the sensory-sequence and motor-sequence predictions for the learned hypersets, that is, extensively experienced sequences. As discussed in Section 5.4.1, even though the posterior putamen is generally considered to be involved in the process of retrieval of information from the LTM, the degree of the involvement is questioned for the learned hypersets in the blockade experiment of different parts of the striatum. In addition, the blockade of the middle-posterior caudate did not change the performance of either learned or new hypersets, even though that portion of the caudate constitutes a part of the oculomotor circuit and would seem to be involved in producing anticipatory saccades. Thus, after (over)training in the case of the learned hypersets in the 2x5 task, it is likely that information of both sensory-sequence and motor-predictions may be transferred somewhere else to be stored. In the case of new hypersets, the sensory-sequence prediction presumably plays a small role in execution. We tested, therefore, whether the performance of the model decreases in the same way as the monkeys' performance in the pre-SMA and the anterior striatum blockade experiment under the condition in which not the sensory-sequence prediction but only the

Figure 7.11

Figure 7.11: Experimental data in conditions before and after blockade of the anterior striatum, taken from Miyachi et al. [44]: the number of error trials to criterion (10 successful trials) (ordinate) for the control condition and the anterior striatum blockade condition (abscissa) for the new hypersets (solid line) and the learned hypersets (broken line).

Figure 7.12

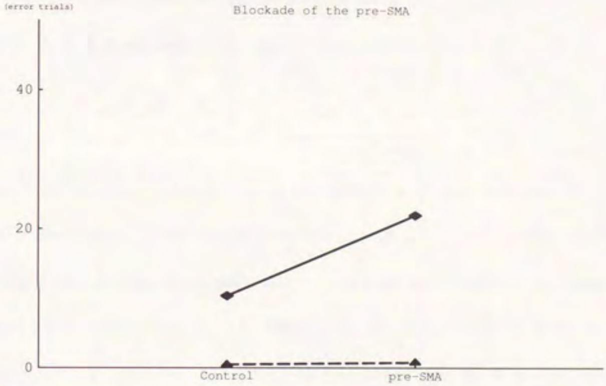


Figure 7.12: Experimental data in conditions before and after the blockade of the pre-SMA, taken from Miyashita et al. [46, 47]: the number of error trials to criterion (10 successful trials) (ordinate) between the control condition and the preSMA blockade condition (abscissa) for the new hypersets (solid line) and the learned hypersets (broken line).

sensory-motor mapping is disturbed.

The blockade of an area by muscimol injection decreases overall neural activity in that area. In the simulation, this effect can be interpreted as a decrease in the output gain of each unit. Recall that the definition of the sigmoid function (equation 6.7) is:

$$F(z) = \frac{1}{1 + exp(-z)}$$

To achieve the effect discussed above as decreasing the output gain, the factor, called *temperature*, T, is introduced to the sigmoid function as follows[2]:

$$F(z, T) = \frac{1}{1 + exp(-\frac{1}{T}z)} \tag{7.1}$$

Figure 7.13 shows the relationship of the output with the input at different temperatures. It is clear that, as the temperature increases, the curve of the sigmoid function becomes less steep and, consequently, the output gain decreases for the same magnitude of the input or becomes closer to 0.5. Recall that the probability of pushing the smaller numbered button decided by the sensory-motor mapping is given in Equation (6.8), using the sigmoid function. To understand the effect of the temperature on this equation, suppose the extreme such that the output is always 0.5, whatever the input is. This means that the probability of pushing either the smaller numbered button or the larger numbered button is always 0.5, that is, equally random, whatever set is given as input. Thus, increasing the temperature results in decreasing the ability to classify inputs and consequently to produce the corresponding outputs. In this way, it is possible to disturb the function of the sensory-motor mapping.

With five different temperatures ($T = 1, 3, 5, 10, 20$), we tested the performance of the learned hypersets and new hypersets, using the weights of the sensory-sequence

---

[2] The name of this factor, temperature, originates from the fact that there exists to some degree an analogy between some aspects of neural networks and statistical mechanics. [24]

Figure 7.13

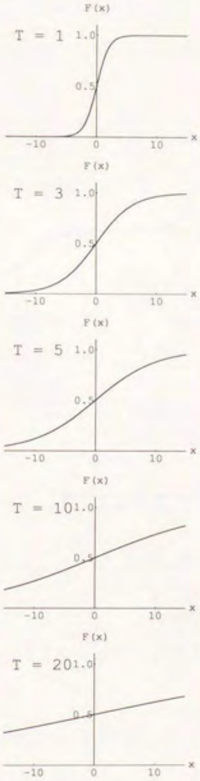


Figure 7.13: Sigmoid function at different temperatures. As the temperature (T) increases, the curve of the sigmoid function becomes less steep. From top to bottom, $T = 1$, $T = 3$, $T = 5$, $T = 10$, and $T = 20$.

Figure 7.14

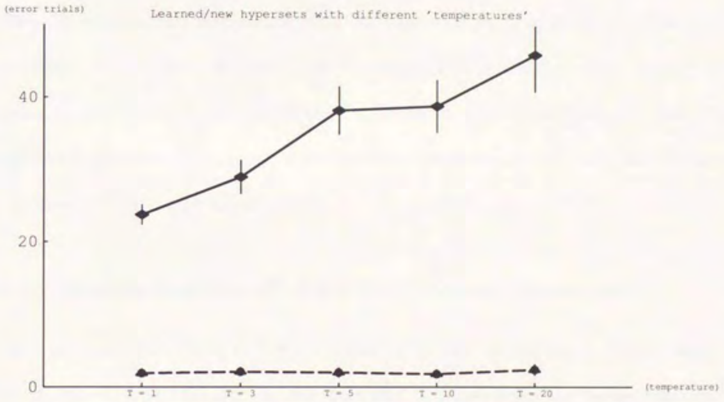


Figure 7.14: Averaged number of error trials to criterion (10 successful trials) (ordinate) obtained by the simulation of the model for the new hypersets (solid line) and learned hypersets (broken line) at different temperatures (abscissa). From left to right, $T = 1$, $T = 3$, $T = 5$, $T = 10$, and $T = 20$. The error bar at each data point stands for the standard error (SE).

prediction and motor-sequence prediction obtained after the training. Note that the condition of the temperature, $T = 1$, is the control condition. The results are shown in Figure 7.14. As the temperature increases, that is, as the output gain decreases, the performance for the new hypersets gets worse. In contrast, the performance for the learned hypersets does not differ greatly between different temperatures. In comparison to the control condition ($T = 1$), the difference in performance for the new hypersets is statistically significantly worse even with the temperature, $T = 3$ ($p < 0.05$) and $T = 5$ ($p < 0.005$). It is, therefore, clear that decreasing the capability of the sensory-motor mapping results in interference with the performance not for the learned hypersets but for the new hypersets. This result captures the experimental results for the blockade of the anterior striatum and the pre-SMA.

7.3.4 Generalization of skills for learned hypersets

It often happens that parts of well-acquired skills can be applied to learn other new skills, as stated in Section 3.1. In this way, the acquired skills can be generalized. We asked whether the model can make use of the acquired ability of learned hypersets to do new hypersets with a specific condition explained below.

We generated the hypersets, called *the modified-learned hypersets*, from the learned hypersets by the following procedures. First, for each learned hyperset, one set is randomly chosen from the 2nd to the 4th sets. Secondly, the chosen set is changed to a new set different from the original set while obeying the constraints of the procedure to generate hypersets discussed in Section 7.2.3. Thus, each of the modified-learned hypersets has the same four sets in the same order as one of the learned hypersets but the other one set is different from that of the learned hyperset. In the above modification procedure, the 1st set is not changed. It is because we need a special consideration on the context for the 1st set, as discussed in Section 6.4.6.

Figure 7.15 shows the comparison of the performance among the learned hypersets, the modified-learned hypersets, and the new hypersets. The average numbers of error trials to criterion (10 successful trials) are plotted. The performance of the learned hypersets (mean 2.4) is significantly better than that of the modified-learned hypersets (mean 14.7, $p < 0.005$) and of the new hypersets (23.8, $p < 0.001$). The performance of the modified-learned hypersets is significantly better than that of the new hypersets ($p < 0.05$).

An example of learning the modified-learned hypersets is seen in Figure 7.16. In this example, the 3rd set is changed from the learned hyperset (No.2) (See Figure 7.16 top left and bottom left). Comparing the change in the number of the completed sets between the modified-learned hyperset (Figure 7.16, top right) and the learned hyperset (Figure 7.16, bottom right), it is clear that the model made most of mistakes at the 3rd, 4th and 5th sets for the modified-learned hyperset. It is because the sensory-motor mapping decides an action for the 3rd set by use of the information of the 3rd set as sensory inputs and the sensory-sequence prediction and motor-sequence prediction *anticipate* the coming sensory inputs and corresponding actions for the 4th and 5th set, partially based on the information of the 3rd set. For other sets, the acquired skills of the learned hyperset (No.2) can be applied so that there are only a few mistakes for these sets. It can be concluded, thus, that the model exhibits the capability of applying parts of the acquired skills (learned hypersets) to learning new skills (modified-learned hypersets). This condition has not been tested yet experimentally in the 2x5 task. It will be interesting to see how the performance of the monkeys resembles or differspu from that of the model.

Figure 7.15

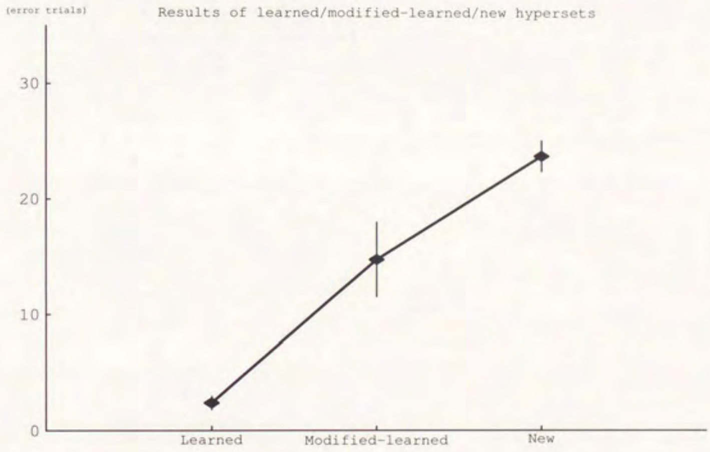


Figure 7.15: Averaged number of error trials to criterion(10 successful trials) of the model for the learned hypersets(left), the modified-learned hypersets (middle), and the new hypersets (right). The error bar at each data point stands for the standard error (SE).

Figure 7.16

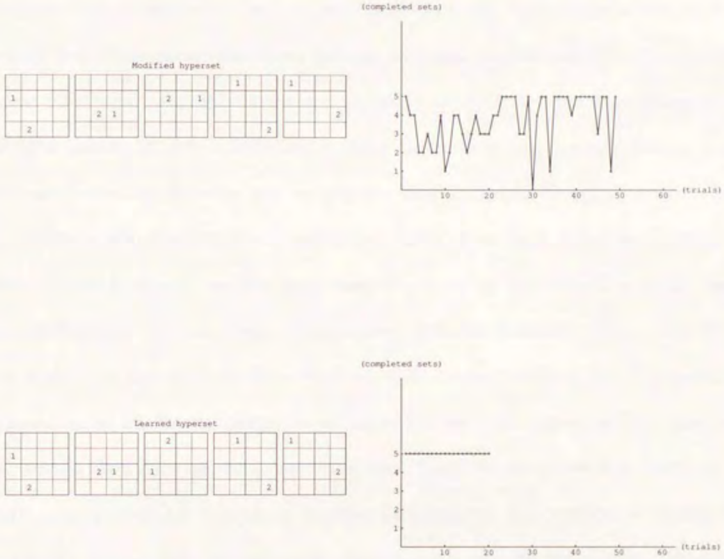


Figure 7.16: Comparison of the performance for the modified-learned hyperset (top, left) and the learned hyperset (No. 2) (bottom, left). The change in the number of completed sets (ordinate) across trials (abscissa) is shown for the modified-learned hyperset (top, right) and for the learned hyperset (bottom, right).

7.4 Discussion

The basal ganglia has been known as involved in the acquisition and execution of sequential movements by experimental studies and empirical studies of brain lesions and diseases. Because of the striking characteristic of the basal ganglia such that it receives from almost the entire cortex and projects back primarily to the frontal cortex, it is conceived that there are several loops linking the basal ganglia and the cerebral cortical areas (the basal ganglia-thalamocortical loops) and that the basal ganglia may be involved in coordinating functions among these loops. It is not, however, known much about how such a coordination can be achieved from a computational viewpoint. This study addressed this computational question in terms of sequential decision making.

Several computational models have been proposed for the functions of the basal ganglia and related cortical areas, in some cases, with the functions of the cerebellum. These models are based on the framework of reinforcement learning (RL), inspired by the experimental results on the profile of neural activities of dopamine (DA) neurons. These models shed light on some aspects of their functions as reviewed in Section 4.7. Though some models are interesting because they address the question of linking the basal ganglia, the cerebellum, and motor cortical areas, none of them has yet addressed the question of the relationship of the loops of the basal ganglia in terms of sequential decision making. None of these researches has done the close examination of the performance of their models in comparison with the actual behavioral data, particularly in terms of sequential decision making.

With their experimental findings in the 2x5 task, Hikosaka et al. [26] posited that there are schematically two functional components for functions of the basal ganglia and related cortical areas: temporary storage and permanent storage. The former is assumed primarily to play a role in the acquisition at the early stage of experiencing the sequences (the learning mechanism in their term), whereas the latter is assumed to play a role in execution in the late stage (the memory mechanism in their term). Their scheme fits to some degree with their experimental findings, but it has not been computationally clear how their two functional components are integrated in both of the early and late stages nor how the learned result in the learning mechanism can be transferred to the memory mechanism.

The computational elements of functions of the loops of the basal ganglia in terms of sequential decision making (sequential movements) and a general framework on the acquisition and retrieval processes in execution are proposed in Chapter 6. A model implementing an algorithm based on the general framework is also given in Chapter 6. It is postulated in this study that (1) the presupplementary motor area (pre-SMA) and anterior striatum function to associate sensory input with its optimal output in the sequences (the sensory-motor mapping); (2) the dorsolateral prefrontal circuit, in particular, the dorsolateral prefrontal cortex and the anterior striatum, functions to learn the sensory-dependent sequences (the sensory-sequence prediction); (3) the motor circuit, in particular, the supplementary motor area (SMA) and the posterior striatum (the posterior putamen), functions to learn the motor-dependent sequences (the motor-sequence prediction). The sensory-sequence prediction is postulated to play a role in the integration of acquisition and retrieval processes in execution. As discussed in Section 6.1, the term, 'context', has been somewhat abused in the experimental and computational literatures as discussed before. In contrast, the term, 'context', is explicitly treated in the model by use of the sensory-sequence prediction and the concept of biasing.

It is shown by the simulation of the model based on the above hypotheses that the qualitative behavior of the model coincides remarkably well with that of the actual monkeys in several aspects: two levels of the learning, that is, the short-term and sequence-selective, and long-term and sequence-selective levels, and the context dependency. It is also shown that when the sensory-motor mapping is blocked in the simulation, the performance of the model resembles well that of the actual monkeys in blockade of the pre-SMA or the anterior striatum. As the output gain of the sensory-motor mapping is more disturbed, the performance of the new hypersets gets worse, whereas that of the learned hypersets almost remains the same. It can be concluded, thus, that the proposed computational elements of functions in the basal ganglia and related cortical areas capture well those of the actual monkeys.

In addition, the performance of the model is tested under condition of the hypersets that have the same four set as one of the learned hypersets but the other set is different from the learned one (the modified-learned hypersets). The model exhibits the capability of applying parts of the acquired sequences of the learned hyperset to learning the modified-learned hypersets. Because this condition is not tested yet experimentally, it should be interesting to examine how the performance of the actual monkeys would resemble or differ from that of the model.

The performance of the model exhibited interesting characteristics of skills at the behavior level, which are discussed in Section 3.1: a multiple time scales of the improvement of skills, context dependency, generalization of skills because of their hierarchical nature, continuity of the development of skills between acquisition and retrieval in execution, and a type of learning, that is, reinforcement learning. Thus, the model captures well the important nature of skills.

In this study, the functions of the basal ganglia and related cortical areas are investigated in terms of skills, one of the long-term memory (LTM) systems in one classification as discussed in the beginning of Chapter 3. It is not, however, asked what function among investigated ones belong to the LTM or not, because of the difficulty of correspondence between behavior and underlying neural mechanism in classifications of the LTM. Rather, this study investigated how the functions of the basal ganglia loops can contribute to the acquisition and retrieval processes in execution of skills. The motor-sequence prediction is presumed to be particularly important for the LTM, as the findings of the 2x5 task measured by the performance time indicates. The sensory-sequence prediction may contribute to information in the LTM as well. Experimental results in the 2x5 task indicates that skills as complex as that of the 2x5 task acquired by these functions can be transferred further to somewhere else to be stored as information in the LTM. Investigating this transfer is one of the future works.

For the simulation in this chapter, there are made several simplifications for the proposed model in order to pinpoint the core of the proposed computational elements. Limitations caused by such simplifications will be discussed in the next chapter, along with the future works.

Chapter 8

CONCLUSION

8.1 Findings

In this thesis, the first study in Chapter 2 focused on the dynamical characteristic of the short-term memory (STM), or working memory, in sequential decision making. The second study from Chapter 3 to Chapter 7 focused on the functions of the basal ganglia and related cortical areas for a particular sequential decision making, or skills that are classified as one of the long-term memories (LTM) in one classification, in relation to reinforcement learning.

In Chapter 2, because neural activities are required for both of loading and maintaining (storing) in working memory, it is asked what type of dynamics of neural activities is suitable for the sequential decision making. Several researchers had proposed that the fixed point attractors enables working memory to robustly store information as neural activities. However, considerations were not given much in the case of the sequential decision making. It is postulated in this study that the long-term maintenance and quick transition of neural activities is a crucial requirement in sequential decision making. We analyzed, first, the dynamics of a a single sigmoidal function that has a self-connection and bias, and, then, the dynamics of a recurrent network that has the uniform self-connections and bias. It was mathematically shown that the specific parameter region, that is, the near saddle-node bifurcation region, can realize the hypothesized crucial dynamics for sequential decision making. The plausibility of this dynamics for sequential decision making is examined in the simulation of foraging tasks. It was shown by evolutionary programming that the near saddle-node bifurcation dynamics emerged to improve the performance in tasks as the environment became severe. Hence, it can be concluded that the proposed dynamics in this study, the near saddle-node bifurcation dynamics, can be a candidate of the dynamics in working memory for sequential decision making, in particular in severe non-stationary environments.

Chapter 3, 4, 5, 6, and 7 are devoted to study the functions of the basal ganglia-thalamocortical loops in sequential decision making. Chapter 3, 4, and 5 provided the reviews and discussions of the basal ganglia, the framework of reinforcement learning, and the behavioral and neurophysiological findings in relation to the basal ganglia loops, respectively. Based on the investigation of these chapters, it is postulated in Chapter 6 that the dorsolateral prefrontal circuit and the motor circuit gradually acquire the sensory-dependent and motor-dependent sequences respectively, that is, the sensory-sequence and motor-sequence predictions, while the presupplementary motor area (pre-SMA) and the anterior striatum map the sensory inputs with the motor outputs (the sensory-motor mapping). The dopamine (DA) neurons in the substantial nigra pars compacta (SNc) are assumed to play a role of providing reinforcement signals for all these three learning processes. A general framework on the acquisition and retrieval processes in execution is, then, provided. Consequently, a model implementing an algorithm based on this framework is given in order to examine the hypothesis in close comparision of the serial button press task (the 2x5 task) developed by Hikosaka and his colleagues.

In Chapter 7, the performance of the model in simulations is compared with several experimental findings in the task. It is shown that the qualitative behaviors of the model replicate these of the monkeys in the experiments in several aspects. It is also shown that when the part of the model that should correspond to the pre-SMA and the anterior striatum in the monkeys are disturbed, the change in the performance of the model resembles the change in the performance of the monkeys. These results may be considered to support our hypothesis of functions in the basal ganglia and related cortical areas both at the behavior and neurophysiological levels. The model exhibited the capability of applying the acquired information to new tasks (modified-learned hypersets), whose condition is not yet tested in the actual experiment of the 2x5 task. The comparison of the simulated results with the experimental results is a future work and may contribute to the direction of future experiments.

8.2 Limitations and future works

Limitations and future works of the studies in this thesis can be listed as follows:

1. **Multiple working memory systems? central controller?**

   It has been argued in the psychological literature that there may be several functional components in working memory. Baddeley [5], for example, proposed 'central executive' that controls the selection and manipulation of information to store in working memory and 'slave system' that is a storage of information in working memory. In the present investigation of working memory, these two functions, manipulating and storing information, are treated as in one system, though the function of manipulation is simplified as only selecting information to store. We investigate how these two functions, which are opposing demands to each other from a view of dynamical systems, can be realized in a single system. It is shown that even without such a central controller, the proposed dynamics of this study serves well as working memory in non-stationary environment. As the tasks become more complicated, a controller may be needed. To address this question is one of future works.

In this study, it is not asked whether there is only one single working memory system or there are multiple working memory systems, for example, visuospatial one and phonological one, each of which can be considered as one of 'slave systems' [5]. As the sensory information becomes more rich and various, multiple working memory systems may be needed other than one single system. In addition, it is often the case that biological systems need to communicate to each other in sequential decision making, for example, a group of people for hunting. It is interesting to see what dynamics would emerge in neural network of the creatures under conditions of multimodal sensory information and/or other creatures. A future work should address this question.

2. **Refinement of the model of functions of the basal ganglia loops**

In the second study of functions of the basal ganglia loops in sequential decision making, the function of the oculomotor circuit is neglected in the simulation of the 2x5 task, even though the functional relationship between the dorsolateral prefrontal and oculomotor circuits were briefly discussed. It is important to integrate the functions of oculomotor circuit with the currently proposed scheme. There are several experimental findings in the 2x5 task regarding the oculomotor circuit such as anticipatory saccade. The performance of the integrated model should be compared with those experimental findings. In the simulation of the 2x5 task, the representations of the sensory inputs and motor outputs were much simplified. Using more realistic representations, for example, representations including the spatial relationship should be needed for futher investigation. Such representations would help to develop more refined model that includes more of the motor control aspects, which is discussed below.

3. **Motor control aspects: integration of functions of the basal ganglia and the cerebellum**

   The aspects of motor control is much simplified in the current simulation of this thesis. Consequently, the functions of the cerebellum and other motor cortical areas such as the primary motor area are not considered in detail. It is important to synthesize the proposed current scheme of functions of the basal ganglia and related cortical areas with those portions of the brain. In this aspect, it would be a good starting point to investigate a way of in-cooperating the proposed scheme with the hypothesis proposed by Houk and Wise [30,31,80] at the theoretical level and with that proposed by Doya [14,15] at the computational level. In experimental findings of the 2x5 task, there are already available several experimental findings for the aspects of motor control, for example, results measured by the performance time and results obtained after the blockade of the dentate nucleus in the cerebellum [37, 38].

   It will be very interesting to examine further whether the extension of the current scheme including the motor control aspects would fit well with those experimental results. This extension for the aspects of motor control can contribute to a better understanding of functions of the brain, including the basal ganglia, the cerebellum, and other cortical areas, in sequential decision making, particularly in sequential movements.

4. **Integration with other neural mechanisms**

Though the second study of the thesis focused functions of the basal ganglia and related cortical areas in relation to skills, other neural mechanisms, corresponding to other types of the LTM, cannot be completely neglected even in the learning of skills. For example, the declarative memory system, one of the LTM, may contribute to the acquisition of sequences particularly in the early stage. The extension of the current scheme with other neural mechanisms or other types of LTM is another direction of future work.

Several researchers in machine learning (and reinforcement learning) have investigated the synthesis of a kind of gradual, iterative-based, and statistical learning with a kind of more quick and declarative learning. For example, Thrun [76] investigated to integrate reinforcement learning (RL) with explanation-based learning, which is more rule-based and, thus, similar to the declarative memory to some degree. This kind of synthesis of the proposed scheme with other memory systems should be interesting from a computational viewpoint.

From a neuroscientific viewpoint, several portions of the brain can be interesting for such synthesis. Amygdala and hippocampus are of particular interest. Amygdala is known as related to associating sensory inputs with emotions and can be also a good candidate to provide primary reinforcement signals to dopamine neurons. The hippocampus has been considered as associating sensory inputs [9] and as being involved in the formation of the declarative memory, in contrast to the striatum that is considered as more gradual and incremental in learning [35]. The synthesis of the current scheme on the functions of the basal ganglia with these portions of the brain will be a good future work.

Appendix A

Range of initial parameters

The range of initial parameters were set as follows: $[5.0, 10.0]$ for a, $[-8.0, -2.0]$ for b, $[-2.0, 0.50]$ for c_1, c_2, c_3, $[-1.5, 0.50]$ for c_4, c_5, $[1.0, 6.0]$ for d, $[-1.0, 0.0]$ for e, $[3.0, 6.0]$ for f_1, f_3, $[5.0, 9.0]$ for f_2, f_4.

Appendix B

Abbreviation

CD = caudate nucleus

DA = dopamine neuron

DLPF = dorsolateral prefrontal cortex

FEF = frontal eye field

fMRI = functional Magnetic Resonance Imaging

GP = globus pallidus

GPi = the internal segment of globus pallidus

GPe = the external segment of globus pallidus

LTM = long-term memory

M1 = primary Motor Cortex

PMC = premotor Cortex

pre-SMA = presupplementary Motor Cortex

Pt = putamen

rCMA = rostral cingulate motor area

RL = reinforcement learning

S1 = primary somatosensory cortex

SC = superior colliculus

SEF = supplementary eye field

SMA = supplementary Motor Cortex

SN = substantia nigra

SNc = substantia nigra pars compacta

SNr = substantia nigra pars reticulata

STN = subthalamic nuclues

STM = short-term memory

Appendix C

Parameters for the model of the basal ganglia and related cortical areas

Learning parameters

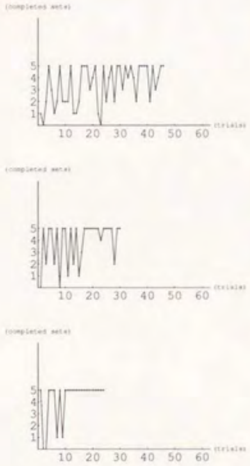
learning parameters for the critic(value function), the sensory-motor mapping, the sensory-sequence prediction, the motor-sequence prediction, and the special weight for the first set are set as 0.4, 0.8, 0.75,0.5, 0.05, respectively.

Appendix D

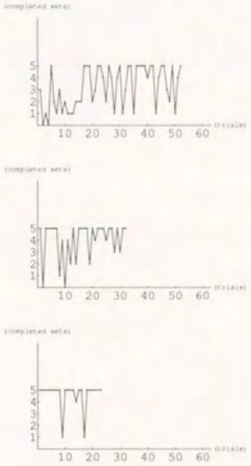
Complementary figures for the simulation of the 2x5 task

From the next page, complementary figures are shown for the simulation of the 2x5 task. Figures are given for all 10 learned hypersets, each of which is numbered on the top. See Figure 7.1 for the information of numbered learned hypersets. See the caption of Figure 7.3 and of Figure 7.4 for the information of the format of figures.

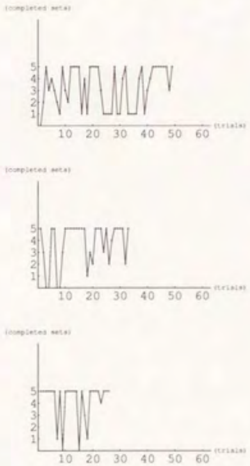
No. 1



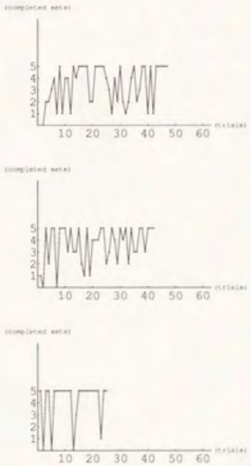
No. 2



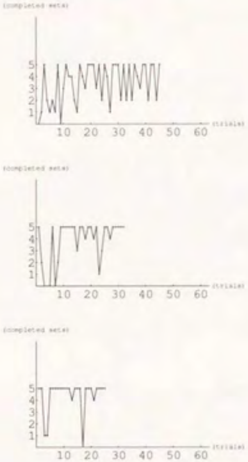
No. 3



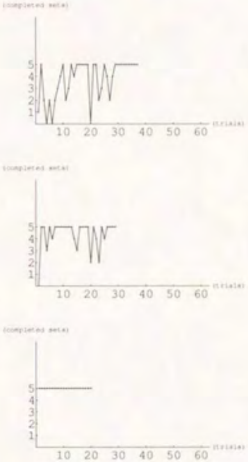
No. 4



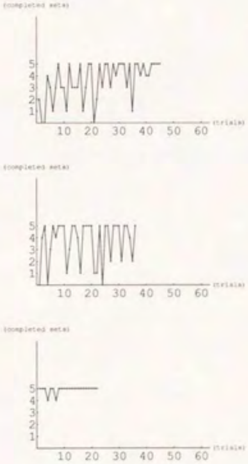
No. 5



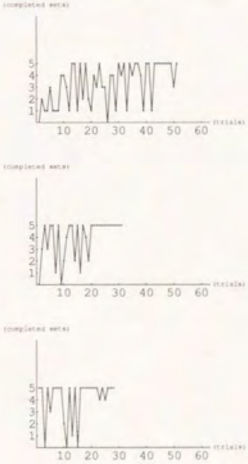
No. 6



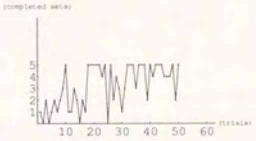
No. 7

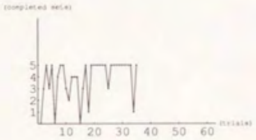


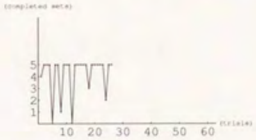
No. 8



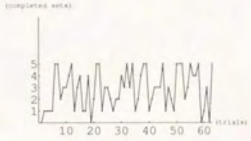
No. 9

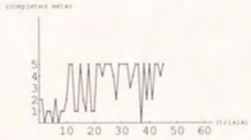

(completed sets)
5
4
3
2
1
10
20
30
40
50
60
(trials)

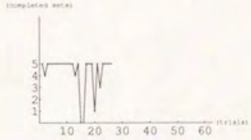

(completed sets)
5
4
3
2
1
10
20
30
40
50
60
(trials)

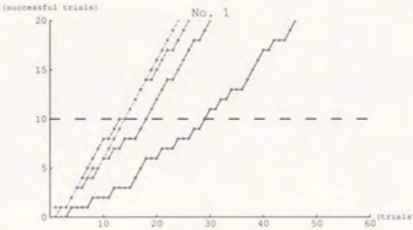

(completed sets)
5
4
3
2
1
10
20
30
40
50
60
(trials)

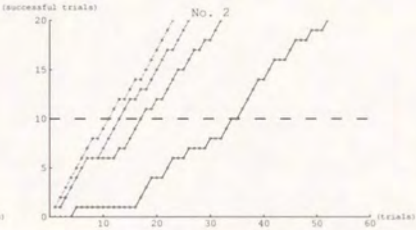
No. 10

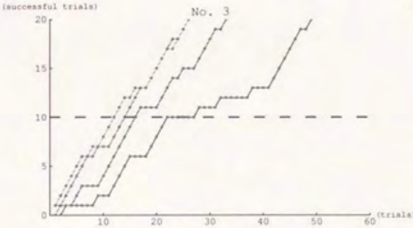

(completed sets)
5
4
3
2
1
10
20
30
40
50
60
(trials)

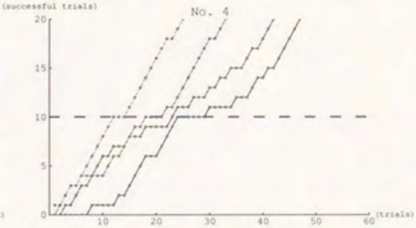

(completed sets)
5
4
3
2
1
10
20
30
40
50
60
(trials)

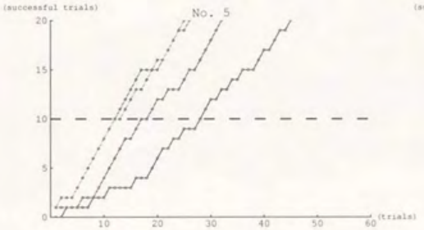

(completed sets)
5
4
3
2
1
10
20
30
40
50
60
(trials)

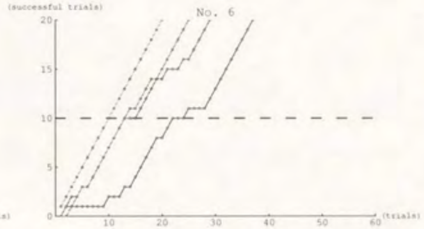

No. 1
(successful trials)
20
15
10
5
0
10
20
30
40
50
60
(trials)

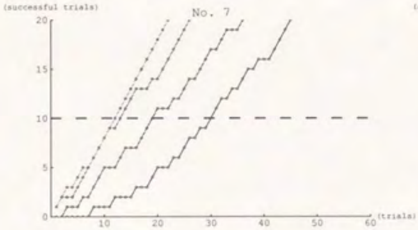

No. 2
(successful trials)
20
15
10
5
0
10
20
30
40
50
60
(trials)

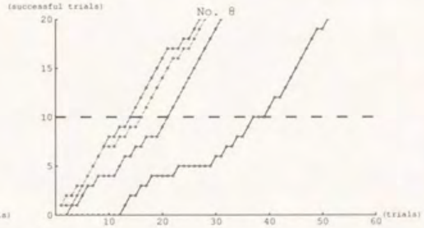

No. 3
(successful trials)
20
15
10
5
0
10
20
30
40
50
60
(trials)

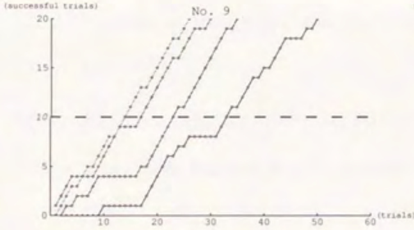

No. 4
(successful trials)
20
15
10
5
0
10
20
30
40
50
60
(trials)

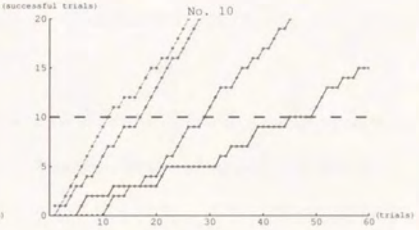

(successful trials)
No. 5
20
15
10
5
0
10
20
30
40
50
60
(trials)


(successful trials)
No. 6
20
15
10
5
0
10
20
30
40
50
60
(trials)


(successful trials)
No. 7
20
15
10
5
0
10
20
30
40
50
60
(trials)


(successful trials)
No. 8
20
15
10
5
0
10
20
30
40
50
60
(trials)


(successful trials)
No. 9
20
15
10
5
0
10
20
30
40
50
60
(trials)


(successful trials)
No. 10
20
15
10
5
0
10
20
30
40
50
60
(trials)

Bibliography

[1] H. Aizawa, M. Inase, H. Mushiake, K. Shima, and J. Tanji. Reoranization of activity in the supplementary motor area associated with motor learning and functional recovery. *Experimental Brain Research*, 84:668–671, 1991.

[2] J. S. Albus. A theory of cerebellar function. *Mathematical Biosciences*, 10:25–61, 1971.

[3] G. Alexander and M. Crutcher. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences*, 13(7):266–271, 1990.

[4] G. E. Alexander, M. D. Crutcher, and M. R. DeLong. Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. In H. Uylings, C. Van Eden, J. De Bruin, M. Corner, and M. Feenstra, editors, *Progress in Brain Research*, volume 85, chapter 6, pages 119–146. Elsevier Science Publishers B.V., 1990.

[5] A. D. Baddeley. Working memory. *Science*, 255:556–559, 1992.

[6] A. Barto, R. Sutton, and C. Anderson. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:834–846, 1983.

[7] A. G. Barto. Adaptive critics and the basal ganglia. In *Models of Information Processing in the Basal Ganglia*, pages 215–232. MIT Press, Cambridge, Massachusetts, 1995.

[8] J. F. Bates and P. Goldman-Rakic. Prefrontal connections of medial motor areas in the rhesus monkey. *Journal of Comparative Neurology*, 336:211–228, 1993.

[9] A. Bechara, D. Tranel, H. Damasio, R. Adolphs, C. Rockland, and A. R. Damasio. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, 269:1115–1118, 1995.

[10] G. Berns and T. Sejnowski. How the basal ganglia make decisions. In A. Damasio, H. Damasio, and Y. Christen, editors, *Neurobiology of Decision-Making*, pages 101–111. Springer-Verlag, Heidelberg: Berlin, 1996.

[11] D. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont: MA, 1995.

[12] L. Chen and S. Wise. Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. *Journal of Neurophysiology*, 73(3):1101–21, Mar 1995.

[13] L. Chen and S. Wise. Supplementary eye field contrasted with the frontal eye field during acquisition of conditional oculomotor associations. *Journal of Neurophysiology*, 73(3):1122–34, Mar 1995.

[14] K. Doya. An integrated model of basal ganglia and cerebellum in sequential control tasks. *Society for Neuroscience Abstract*, 22:2029, 1996.

[15] K. Doya. Efficient nonlinear control with actor-tutor architecture. In D. Touretzky, M. Mozer, and M. Hasselmo, editors, *Advances in Neural Information Processing Systems*, volume 9, 1997.

[16] D. Fogel, L. Fogel, and V. Porto. Evolving neural networks. *Biological Cybernetics*, 63:487–493, 1990.

[17] S. Funahashi, M. Inoue, and K. Kubota. Delay-related activity in the primate prefrontal cortex during sequential reaching tasks with delay. *Neuroscience Research*, 18:171–175, 1993.

[18] S. Funahashi and K. Kubota. Working memory and prefrontal cortex. *Neuroscience Research*, 21:1–11, 1994.

[19] P. S. Goldman-Rakic. Working memory dysfunction in schizophrenia. *Journal of Neuropsychiatry*, 6:348–357, 1994.

[20] J. Guckenheimer and P. Homes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields.* Springer-Verlag, 1983.

[21] U. Halsband, N. Ito, J. Tanji, and H. J. Freund. The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain*, 116:243–266, 1993.

[22] U. Halsband, Y. Matsuzaka, and J. Tanji. Neuronal activity in the primate supplementary, pre-supplementary and premotor cortex during externally and internally instructed sequential movements. *Neuroscience Research*, 20(2):149–55, 1994.

[23] S. Haykin. *Neural Networks, A Comprehensive Foundation.* Macmillan, New York, NY, 1994.

[24] J. Hertz, A. Krogh, and R. Palmer. *Introduction to the Theory of Neural Computation.* Addison-Wesley Publishing Company, Inc., Redwood City, CA, 1991.

[25] O. Hikosaka, M. Kato, S. Miyachi, and K. Miyashita. Learning of sequential movements in the monkey: Process of learning and retention of memory. *Journal of*

Neurophysiology, 74(4):1652–1661, 1995.

[26] O. Hikosaka, S. Miyauchi, K. Miyashita, and M. K. Rand. Learning of sequential procedures in monkeys. In J. Bloedel, T. Ebner, and S. Wise, editors, *The Acquisition of Motor Behavior in Vertebrates*, pages 303–317. MIT Press, Cambridge, MA, 1996.

[27] O. Hikosaka, M. K. Rand, S. Miyauchi, and K. Miyashita. Procedural learning in the monkey. In M. Kimura and A. Graybiel, editors, *Functions of the Cortico-Basal Ganglia Loop*, pages 18–30. Springer, Tokyo, 1995.

[28] O. Hikosaka, K. Sakai, T. Miyachi, R. Takino, Y. Sasaki, and B. Puetz. Human pre-SMA and precuneus. In *Proceedings of 19th Annual Meeting of The Japan Neuroscience Society.*, page 2316, 1996.

[29] O. Hikosaka, K. Sakai, S. Miyauchi, R. Takino, Y. Sasaki, and B. Putz. Activation of human pre-SMA in learning of sequential procedures – a functional MRI study. *Journal of Neurophysiology*, 76(1):617–621, 1996.

[30] J. Houk and S. Wise. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex*, 5:95–110, 1995.

[31] J. C. Houk, J. L. Adams, and A. G. Barto. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of Information Processing in the Basal Ganglia*, pages 249–270. MIT Press, Cambridge, Massachusetts, 1995.

[32] L. Kaelbling, M. Littman, and A. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelliegence Research*, 4:237–285, 1996.

[33] R. Kandel, J. H. Schwartz, and T. M. Jessell. *Principles of Neural Science*. Appleton & Lange, 1991.

[34] M. Kawato. Optimization and learning in neural networks for formation and control of coordinated movement. In D. Meyer, editor, *Attention and Performance*, volume XIV, pages 1–40. Lawrence Erlbaum, Hillsdale, New Jersey, 1991.

[35] B. Knowlton, J. Mangels, and L. Squire. A neostriatal habit learning system in humans. *Science*, 273:1399–1402, Sep 1996.

[36] P. H. Lindsay and D. A. Norman. *Human Information Processing: An Introduction to Psychology*. Academic Press, 2nd edition edition, 1977.

[37] X. Lu, O. Hikosaka, and S. Miyachi. Role of monkey cerebellar dentate nucleus in procedural memory. In *Proceedings of 26th Annual Meeting, Society for Neuroscience.*, page 546.13, 1996.

[38] X. Lu, O. Hikosaka, and S. Miyauchi. Monkey cerebellum. In *Proceedings of 19th Annual Meeting of The Japan Neuroscience Society*, page 2317, 1996.

[39] G. Luppino, M. Matelli, R. Camarda, and G. Rizzolatti. Corticocortical connections of area f3 (SMA-proper) and are f6 (pre-SMA) in the Macaque monkey. *The Journal of Comparative Neurology*, 338:114–140, 1993.

[40] Y. Matsuzaka, H. Aizawa, and J. Tanji. A motor area rostral to the supplementary motor area (presupplementary motor area) in the monkey: neuronal activity during a learned motor task. *Journal of Neurophysiology*, 68(3):653–662, 1992.

[41] J. D. McAuley and J. Stampfli. Analysis of the effects of noise on a model for the neural mechanism of short-term active memory. *Neural Computation*, 6(4):668–678, 1994.

[42] J. Mirenowicz and W. Schultz. Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*, 72(2):1024–1027, 1994.

[43] J. Mirenowicz and W. Schultz. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*, 379:449–451, 1996.

[44] S. Miyachi, O. Hikosaka, K. Miyashita, Z. Karadi, and M. Kato. Different roles of monkey straitum in learning of sequential movements. *Experimental Brain Research*, in press, 1996.

[45] K. Miyashita, M. Kato, S. Miyachi, and O. Hikosaka. Anticipatory saccades in sequential procedual learning in monkeys. *Journal of Neurophysiology*, 76(2):1361–1366, 1996.

[46] K. Miyashita, K. Sakai, and O. Hikosaka. Effects of SMA and pre-SMA inactivation on learning of sequential movements in monkey. In *Proceedings of 26th Annual Meeting, Society for Neuroscience.*, page 731.3, 1996.

[47] K. Miyashita, K. Sakai, and O. Hikosaka. Monkey SMA and preSMA. In *Proceedings of 19th Annual Meeting of The Japan Neuroscience Society.*, page 2315, 1996.

[48] R. Montague, P. Dayan, and T. J. Sejnowski. Framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, 16:1936–1947, 1996.

[49] H. Nakahara and K. Doya. Dynamics of attention as near saddle-node bifurcation behavior. In D. Touretzky, M. Mozer, and M. Hasselmo, editors, *Advances in Neural Information Processing Systems 8.*, pages 38–44, 1996.

[50] S. Nolfi, J. Elman, and D. Parisi. Learning and evolution in neural networks. Technical Report 9019, CRL, University of California, San Diego, 1990.

[51] H. Parthasarathy, J. Schall, and A. Graybiel. Distributed but convergent ordering of corticostriatal projections: Analysis of the frontal eye field and the supplementary eye field in the Macaque monkey. *The Journal of Neuroscience*, 12:4468–4488, 1992.

[52] P. Roland, B. Larsen, N. Lassen, and E. Skinhof. Supplementary motor area and other cortical areas in organizaton of voluntary movement in man. *Journal of Neurophysiology*, 43:118–136, 1980.

[53] R. Romo, E. Scarnati, and W. Schultz. Role of primate basal ganglia and frontal cortex in the internal generation of movements. II. movement-related activity in the anterior striatum. *Experimental Brain Research*, 91(3):385–95, 1992.

[54] J. A. Saint-Cyr, A. E. Taylor, and A. E. Lang. Procedural learning and neostriatal dysfunction in man. *Brain*, 111:941–959, 1988.

[55] K. Sakai and O. Hikosaka. Error correction response in the medial frontal cortex. unpublished manuscript.

[56] K. Sakai, O. Hikosaka, S. Miyauchi, R. Takino, Y. Sasaki, and R. Putz. Activation of human pre-SMA in early acquisition of sequential memory. In *Proceedings of 26th Annual Meeting, Society for Neuroscience.*, page 731.2, 1996.

[57] J. Schmidhuber. Learning to control fast-weight memories: An alternative to dynamic recurrent networks. *Neural Computation*, 4(1):131–139, 1992.

[58] W. Schultz, P. Apicella, and T. Ljungberg. Responses of monkey dopamine neurons to reward and conditio ned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, 13(3):900–913, 1993.

[59] W. Schultz, P. Apicella, R. Romo, and E. Scarnati. Context-dependent activity in primate striatum reflecting past and future behavioral events. In *Models of In-*

formation Processing in the Basal Ganglia, pages 11–27. MIT Press, Cambridge, Massachusetts, 1995.

[60] W. Schultz and R. Romo. Neuronal activity in the monkey striatum during the initiation of movements. *Experimental Brain Research*, 71:431–436, 1988.

[61] W. Schultz and R. Romo. Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology*, 63(3):607–624, March 1990.

[62] W. Schultz and R. Romo. Role of primate basal ganglia and frontal cortex in the internal generation of movements. I. preparatory activity in the anterior striatum. *Experimental Brain Research*, 91(3):363–84, 1992.

[63] W. Schultz, R. Romo, T. Ljungberg, J. Mirenowicz, J. R. Hollerman, and A. Dickinson. Rewared-related signals carried by dopamine neurons. In *Models of Information Processing in the Basal Ganglia*, pages 231–248. MIT Press, Cambridge, Massachusetts, 1995.

[64] T. J. Sejnowski and P. S. Churchland. *The Computational Brain.* MIT Press, Massatusets, 1992.

[65] L. Selemon and P. Goldman-Rakic. Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *The Journal of Neuroscience*, 5(3):776–794, 1985.

[66] R. M. Shiffrin. Short-term memory: A brief commentary. *Memory & Cognition*, 21(2):193–197., March 1993.

[67] K. Shima, H. Mushiake, N. Saito, and J. Tanji. Role for cells in the presupplementary motor area in updating motor plans. *Proceedings of the National Academy of Sciences of the United States of America*, 93(16):8694–8, Aug 6 1996.

[68] L. Squire, B. Knowlton, and G. Musen. The structure and organization of memory. *Annual Review of Psychology*, 44:453–95, 1993.

[69] P. L. Strick, R. P. Dum, and N. Picard. Macro-organization of circuits connecting the basal ganglia with the cortical motor areas. In *Models of Information Processing in the Basal Ganglia*, pages 117–130. MIT Press, Cambridge, Massachusetts, 1995.

[70] R. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44, 1988.

[71] R. Sutton and A. Barto. Time-derivative models of pavlovian reinforcement. In M. Gabriel and J. Moore, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, pages 497–537. MIT Press, Cambridge, MA, 1990.

[72] S. Sutton, Richard. *Temporal Credit Assignment In Reinforcement Learning.* PhD thesis, University of Massachusetts, MA, Feburary 1984.

[73] J. Tanji. The supplementary motor area in the cerebral cortex. *Neuroscience Research*, 19(3):251–68, May 1994.

[74] J. Tanji and K. Shima. Role for supplementary motor area cells in planning several movements ahead. *Nature*, 371:413–416, 1994.

[75] G. Tesauro. TD-gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation*, 6(2):215–219, 1994.

[76] S. Thrun. Learning to play the game of chess. In G. Tesauro, D. Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems*, volume 7, pages 1069–1076. The MIT Press, 1995.

[77] E. Tulving. Concepts of human memory. In L. R. Squire, N. M. Weinberger, G. Lynch, and J. L. McGaugh, editors, *Memory: Organization and locus of change*, pages 3–32. Oxford University Press, New York, NY, 1991.

[78] M. Watanabe. Reward expectancy in primate prefrontal neurons. *Nature*, 382:629–632, 1996.

[79] S. M. Williams and P. S. Goldman-Rakic. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cerebral Cortex*, 3:199–222, 1993.

[80] S. Wise and J. Houk. Modular neuronal architecture for planning and controlling motor behavior. *Biological Communications of the Danish Royal Academy of Sciences*, 43:21–33, 1994.

[81] D. Zipser. Recurrent network model of the neural mechanism of short-term active memory. *Neural Computation*, 3(2):179–193, 1991.

[82] D. Zipser, B. Kehoe, L. Gwen, and J. Fuster. A spiking network model of short-term active memory. *Journal of Neuroscience*, 13(8):3406–3420, August 1993.

Kodak Color Control Patches
Blue
Cyan
Green
Yellow
Red
Magenta
White
3/Color
Black
© Kodak, 2007 TM: Kodak
Kodak Gray Scale
A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19
C Y M
© Kodak, 2007 TM: Kodak