

Multiple Representations and Algorithms for Sequence Learning

Kenji Doya¹, Hiroyuki Nakahara², Raju S. Bapi^{1,3}, and Okihide Hikosaka⁴
doya@erato.atr.co.jp, hiro@brain.riken.go.jp, bapics@uohyd.ernet.in, hikosaka@med.juntendo.ac.jp

¹Kawato Dynamic Brain Project, ERATO, Japan Science and Technology Corporation
2-2 Hikaridai, Seika, Soraku, Kyoto 619-0288, Japan

²Laboratory for Information Synthesis, RIKEN Brain Science Institute, Japan

³Department of Computer/Information Sciences, University of Hyderabad, India

⁴Department of Physiology, School of Medicine, Juntendo University, Japan

Humans and primates can learn a large variety of skills and patterns of movement. How is the information about their spatio-temporal patterns encoded in our brain? Recent animal experiments and human brain imaging studies suggest that the global networks linking specific parts of the cerebellum, the basal ganglia, and the cerebral cortex are involved in learning and execution of sequential movement. Execution of sequential movement initially requires attention but becomes automatic after repeated practice. We postulate that this process is supported by the use of different coordinate systems as well as different algorithms.

In a visuo-motor task, a sequence can be defined either as a series of moving targets in the visual space or as a series of body movement (Figure 1B). We propose that visual representation of sequence is used in the circuit linking the anterior basal ganglia and the prefrontal cortex (the visual network) while body-based representation is used in the circuit linking the posterior basal ganglia and the motor cortex (the motor network). Sequence representation using the visual coordinate is advantageous for quick learning whereas sequence representation in the motor coordinate is advantageous for real-time control.

We built a neural network model of visuo-motor sequence learning which included two recurrent networks, one using visual coordinate and the other using motor coordinate (Figure 2). Both networks were trained by “temporal difference (TD)” learning, which was recently proposed as a functional model of the midbrain dopaminergic system (Schultz et al., 1997). The model was used to simulate a sequence learning task, called the “2x5

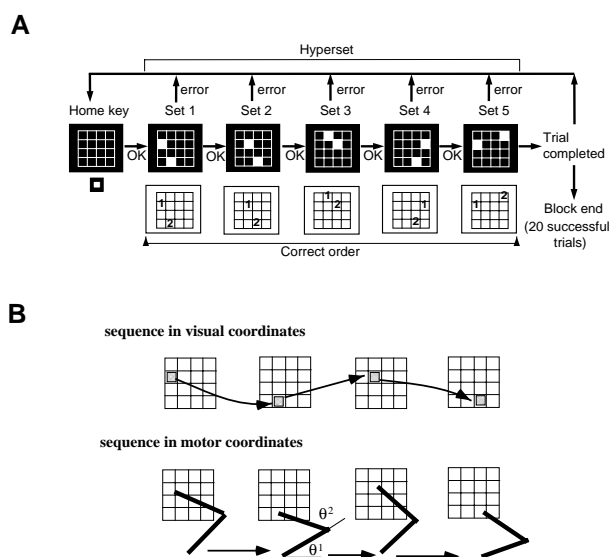


Figure 1: **A.** The 2x5 task for monkeys (Hikosaka et al., 1995). A subject presses five sets two LED buttons in a pre-determined order, which has to be found by trial and error. Liquid reward is given after completion of each *set* of two key presses. The amount of reward is increased as the progress through a five-set sequence, called a *hyperset*. When the subject makes an error, the trial is terminated and restarted. A hyperset is used repeatedly until the number of successfully completed trials reaches to a criterion (10 or 20). Each training day, several hypersets are used for training; some of them are repeatedly used everyday and others are newly generated. **B.** Two possible ways of representing a movement sequence (Nakahara, 1997; Nakahara et al., 1997). Upper panel: to encode the sequence of visuo-spatial locations of the targets to be reached (visual sequence). Lower panel: to encode the sequence of desired postures, for example, using arm joint angles (motor sequence).

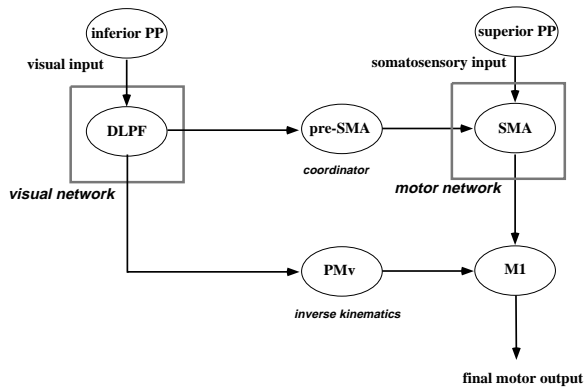


Figure 2: The network model of sequence learning using visual and motor representations (Nakahara et al., 1997). The visual network corresponds to the anterior basal ganglia and its target cortical areas, such as the dorsolateral prefrontal cortex (DLPF). The motor network corresponds to the posterior basal ganglia and its target cortical areas including the supplementary motor area (SMA). The pre-SMA, which links the visual and motor networks, works as a coordinator between the two networks by modulating the output of the motor network based on the output of the visual network. The ventral premotor are (PMv) transforms the visual network output, which is in visuo-spatial coordinate (target spatial position), into motor coordinate (desired joint angles). The outputs of the visual and motor networks converge at the primary motor cortex (MI). A sequence is learned simultaneously by the visual and motor networks based on a reinforcement learning algorithm known as “temporal difference (TD)” learning (Sutton and Barto, 1998).

task” for monkeys (Figure 1A). It was found in reversible lesion experiments that the blockade of the anterior basal ganglia disrupted acquisition of new sequences whereas the blockade of the posterior basal ganglia disrupted execution of well learned sequences (Miyachi et al., 1997). The network model replicated this experimental result with the blockade of the visual and motor networks, respectively (Nakahara, 1997; Nakahara et al., 1997).

We further performed human behavioral experiments to investigate the hypothesis that a subject initially uses visual representation and gradually depends more on motor representation of sequence. Subjects were trained to press a series of keys on a keypad in response to visual stimuli on a display (Figure 3). Response time performance was tested in two altered conditions: VISUAL condition in

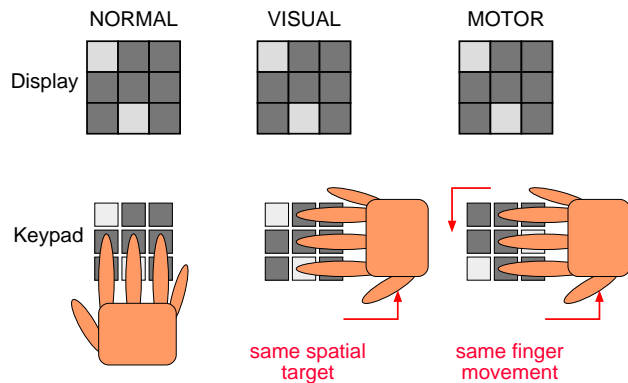


Figure 3: The experiment for assessing the representations for sequence learning (Bapi et al., 1998, 1999). A subject is trained in the 2x10 task, which is similar to the 2x5 task described in Figure 1, in the NORMAL display-keypad condition. The performance of the subject is tested in two altered conditions. In VISUAL condition, the orientation of the hand is rotated 90 degrees, requiring different finger movement for pressing the same key sequence. In MOTOR condition, the keypad is also rotated 90 degrees, resulting in the same finger movement as in NORMAL to reach spatially relocated keys.

which the same key was pressed using different finger movement and MOTOR condition in which the keys were relocated but the finger movement was the same. The response time was initially similar for both conditions, but after about one hour of practice, significantly better transfer was seen in MOTOR condition (Bapi and Doya, 1998; Bapi, et al., 1999).

In addition to the use of multiple representations, the differential involvement of brain areas in early and late stages of sequence learning may also be due to the use of different action selection algorithms. Experimental and theoretical evidence suggests that the cerebellum, the basal ganglia, and the cerebral cortex are specialized, respectively, in supervised, reinforcement, and unsupervised learning paradigms (Doya, 1999). The theory of reinforcement learning and dynamic programming provides several candidate algorithms for utilizing those learning modules.

The simplest algorithm involves stochastic action selection and evaluation of the current state (Figure 4A). The evaluation network learns to predict future reward from the current state. The temporal difference in the

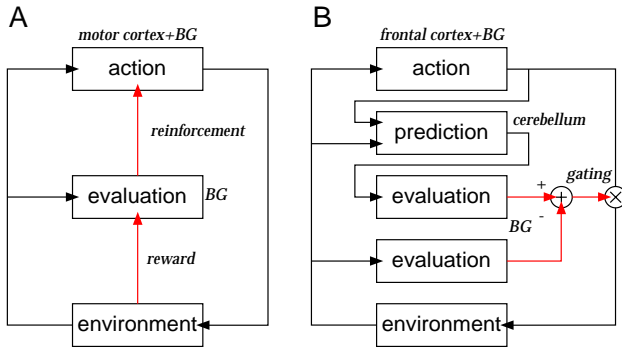


Figure 4: Possible implementation of different sequence learning algorithms in the circuit linking the cerebellum, the basal ganglia (BG), and the cerebral cortex. **A.** A reactive, stochastic action selection network is trained by the reinforcing signal from the evaluation network. The evaluation network trains itself by comparing its prediction with the actual delivery of reward. Such a mechanism can be implemented in the network linking the motor cortex and the posterior basal ganglia. **B.** An internal model of environmental dynamics predicts the future state for a candidate action under the current state. The evaluation of the predicted state is compared with that of the current state. The candidate action is put to execution if it is expected to improve the state. Such a mechanism can be implemented in the network linking the frontal cortex, the lateral cerebellum, and the anterior basal ganglia. The reinforcement signal in **A** as well as the gating signal in **B** are carried by the midbrain dopamine neurons.

predicted reward is used as the reinforcement signal for the stochastic action selection network. In a more elaborate algorithm (Figure 4B), an internal model of the environmental dynamics is used for predicting the future state if a candidate action is taken. The evaluations for the predicted and current states are compared and the action is executed if it is expected to improve the evaluation. Under anatomical and physiological constraints, the reactive, stochastic action selection algorithm (Figure 4A) can be implemented in the network linking the motor cortex and the basal ganglia. The predictive action selection algorithm (Figure 4B) can be implemented in the network linking the frontal cortex, the lateral cerebellum, and the basal ganglia. This hypothesis is in accordance with recent brain imaging studies on sequence learning (Doya, 1999).

In summary, humans and primates can utilize multiple representations and algorithms

for learning and control of sequential movement. Multiple parallel loops linking the cerebellum, the basal ganglia, and the cerebral cortex would enable both quick acquisition and robust execution of motor skills. Similar mechanisms may also be used for cognitive tasks that involve sequential processing.

References

- Bapi, R. S., and Doya, K. (1998) Evidence for effector independent and dependent components in motor sequence learning. *Society for Neuroscience Abstracts*, 24:167.
- Bapi, R. S., Doya, K., and Harner, A. M. (1999) Evidence for effector independent and dependent representations and their differential time course of acquisition during motor sequence learning. *Kawato Dynamic Brain Project Technical Report*, KDB-TR-03.
- Doya, K. (1999) What are the computations of the cerebellum, the basal ganglia, and the cerebral cortex? *Neural Networks*, in press.
- Hikosaka, O., Rand, M. K., Miyachi, S., and Miyashita, K. (1995) Learning of sequential movements in the monkey: Process of learning and retention of memory. *Journal of Neurophysiology*, 74, 1652-1661.
- Miyachi, S., Hikosaka, O., Miyashita, K., Karadi, Z., and Rand, M. K. (1997) Differential roles of monkey striatum in learning of sequential hand movement. *Experimental Brain Research*, 151, 1-5.
- Nakahara, H. (1997) Sequential decision making in biological systems. Ph.D. Thesis, University of Tokyo.
- Nakahara, H., Doya, K., Hikosaka, O., and Nagano, S. (1997) Multiple representations in the basal ganglia loops for acquisition and execution of sequential motor control. *Society for Neuroscience Abstracts*, 23:778.
- Schultz, W., Dayan, P., and Montague, P. R. (1997) A neural substrate of prediction and reward. *Nature*, 275, 1593-1599.
- Sutton, R. S. and Barto, A. G. (1998) *Reinforcement Learning: An Introduction*. MIT Press.