

# Imitating others by composition of primitive actions: a neuro-dynamic model

Hiroaki Arie<sup>a</sup>, Takafui Arakaki<sup>b</sup>, Shigeki Sugano<sup>b</sup>, Jun Tani<sup>a</sup>

<sup>a</sup>*Brain Science Institute, RIKEN*

<sup>b</sup>*Department of Advanced Mechanical Engineering, Waseda Univ.*

---

## Abstract

This paper introduces a novel neuro-dynamical model that accounts for possible mechanisms of action imitation and learning. It is considered that imitation learning require at least two classes of generalization. One is generalization over sensory-motor trajectory variances, and the other class is on cognitive level which concerns on more qualitative understanding of compositional actions by own and others which do not necessarily depend on exact trajectories. This paper describes a possible model dealing with these classes of generalization with focusing on the problem of action compositionality. The model was evaluated in the experiments using a small humanoid robot. The robot was trained with a set of different actions concerning object manipulations which can be decomposed into sequences of action primitives. Then the robot was asked to imitate a novel compositional action demonstrated by human subject which are composed from prior-learned action primitives. The results showed that the novel action can be successfully imitated by decomposing and composing it with the primitives by means of organizing unified intentional representation hosted by mirror neurons even though the trajectory-level appearance is different between the ones of observed and those of self-generated.

---

## 1. Introduction

The understanding of other's actions and imitating them is a significant cognitive capability for humans. Human adults can easily acquire various actions within a very short time by watching other's actions. The learning of such skilled actions should not be considered just as recording a set of experienced motor trajectories in rote memory but as acquiring certain structures

from them. One of such essential structures should be so-called the “behavior compositionality” [1]. The term compositionality here is adopted from “Principle of compositionality” [2] in linguistics. The principle claims that the meaning of a complex expression is determined by the meanings of its constituent expressions and the rules used to combine them. This principle can be adopted also in action domain – skilled actions can be generated by combining a set of re-usable action primitives adaptively by following rules. For example, an attempt of drinking a cup of water might be decomposed into multiple action primitives such as reaching to a cup, grasping the cup and moving the cup toward one’s mouth. Each action primitive can be re-utilized also as a component for other actions e.g. reaching to a cup can be used for another goal such as clearing it away.

This idea of decomposition of whole actions into sequences of reusable primitives has been considered by Arbib [3] as motor schemata theory. The theory considers that a set of action primitives might be preserved in a memory pool and the higher level might retrieve adequate primitives from the pool in order to combine them in generating desired actions. A substantial number of engineering as well as theoretical researches have attempted to implement this idea into autonomous agents including robots [4, 5, 6] as a key mechanism to generate their diverse and complex actions. The typical implementations are to prepare a set of local modules for registering action primitive patterns by using soft computing scheme such as neural network models [7] or fuzzy systems [8] in the lower level and combining those primitives into sequences by manipulating pointers to them symbolically in the higher level. However, such hybrid approaches are likely to have symbol grounding problems [9]. This is because the interface between the higher level of symbol processing and the lower level of pattern processing can be just arbitrary since the symbol systems consisting of arbitrary shapes of tokens cannot share the same metric space with the analog patterns in their interactions [10, 11].

On this account, Tani and his colleagues [10, 11, 12] considered that whole systems might be better constructed on seamless analog dynamical systems instead of having hybrids of symbol and pattern systems. This idea seemed possible because theory of symbolic dynamics [13, 14] had already shown that dynamical systems can exhibit equivalence of symbol-like computation by means of combinatorial mechanics of chaos and fractals. It had been also shown that adaptive dynamical systems, especially recurrent neural network (RNN) model can acquire equivalences of desired symbolic systems of finite

state machines (FSM) by self-organizing fractals [15, 16] or chaos [17] in the internal dynamics. Our expectation has been that cognitive processes of being compositional, yet contextual and situated to the reality, could be resulted by means of the seamless interactions between the top-down cognitive processing and the bottom-up sensory-motor processing in the embodied neuro-cognitive dynamical systems under the name of “organic compositionality” [1].

Now, we explain how dynamical systems framework can be utilized to achieve compositionality in generating diverse actions more specifically. Firstly, we explain simpler problems of how repertory of multiple sensory-motor primitives can be embedded in dynamic neural network through learning. There are two distinct schemes: one is local representation scheme and the other is distributed representation scheme. In the local representation scheme, each sensory-motor pattern is learned as forward model in a specific local dynamic neural network.

Each forward model predicts how sensory-motor state  $x$  changes in time by means of its dynamic function  $f(x, c)$  where  $c$  is the internal state which is not observable from the external. This type of dynamic function can be realized by continuous time recurrent neural network model (CTRNN) [18, 19]. The output of a corresponding local network can be selected externally among others, for example, by utilizing a winner-take-all dynamics (WTA) type gating mechanism [20, 21].

On the other hand in the distributed representation scheme, a set of sensory-motor primitives are embedded in a single dynamic network. How can each distinct repertory of sensory-motor pattern can be selectively generated? There are two ways to achieve this. One way is to utilize the parameter bifurcation characteristics of nonlinear dynamical systems. A dynamic function associated with a parameter  $p$ , which is denoted as  $f(x, c, p)$ , can generate multiple dynamic patterns by modulating the value of  $p$ . Tani et al. [22] implemented this idea in a neural network model so-called the recurrent neural network with parametric biases (RNNPB).

The other way is to utilize the initial sensitivity characteristics of nonlinear dynamical systems of which mechanism will be the focus in the current paper. The idea is schematized in Figure1(a). The dynamic function can generate different trajectories of  $x$  depending on initial internal state  $c_{t=0}$  given. Based on this idea, Nishimoto et al. [23] showed that a humanoid robot learns to generate three different actions of ( $\alpha$ ): grasping object and moving it to left, ( $\beta$ ): grasping object and lifting it up, and ( $\gamma$ ): grasping

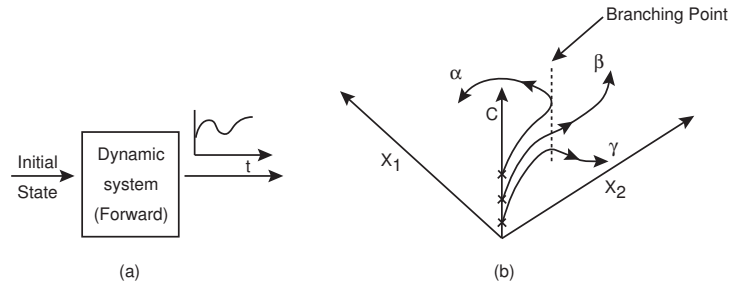


Figure 1: (a) The dynamic function generates different trajectories of  $x$  depending on initial internal state  $c_{t=0}$ . (b) An example of generating three different  $x$  trajectories which trifurcate after the branching point depending on the initial internal state values.

object and moving it to right by starting from the same home position via repeated supervised training of trajectories. It was shown that the differences in the initial internal state values which are self-determined through learning enables the robot trajectories to trifurcate after the robot grasps the object (see the schematics of the idea in Figure1(b)). In this situation, initial internal states is regarded as setting of intentions for particular action programs.

The above mentioned scheme are further extended to be applied to goal-directed planning problems [24]. If distal states of sensory trajectories of learned are taken as goal states e.g., object moved to left, upward or to right in the previous example, the initial internal states which lead to these goal states can be inversely computed (Figure2(a)). This inverse computation re-

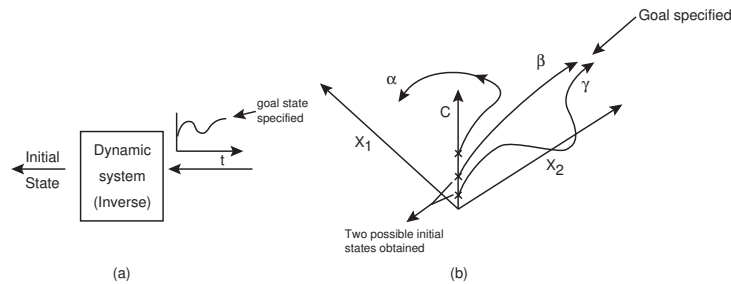


Figure 2: (a) The internal initial state for achieving a distal goal state is inversely computed. (b) Multiple initial states can exist for achieving the same distal goal state.

quires iterative search for finding an optimal initial state value. This process

corresponds to deliberative planning processes. There could be cases that the same distal goal state can be achieved by multiple trajectories as shown in Figure2(b). In such cases, multiple plans in terms of these corresponding initial states should be found through iterative search [24].

It is generally thought that introduction of hierarchy can increase behavioral complexity of the system significantly. Yamashita and Tani [25] showed that functional hierarchical structure responsible for generating action primitives and sequencing of them can emerge in dynamics of CTRNN which consists of multiple time-scale dynamics as illustrated in Figure3. In the

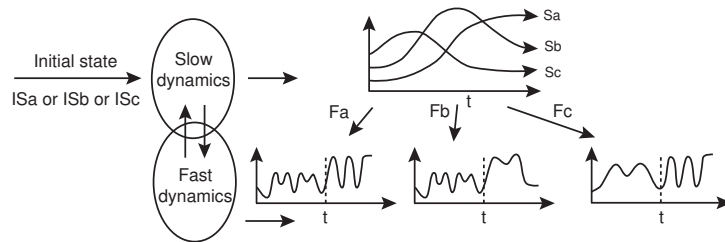


Figure 3: Functional hierarchical structures appear in a dynamic neural network model consisting of multiple time-scales (fast and slow) dynamics.

slow dynamics part different profiles of slowly changing dynamic patterns (Sa, Sb or Sc) appear depending on action programs in terms of the initial state (ISa, ISb or ISc) given which turn out to drive the fast dynamics part to generate corresponding sequences of primitive patterns (Fa, Fb or Fc).

Now let us consider possible correspondences of above mentioned models to system level neuroscience. Tani et al. [1] speculated that setting or determining initial internal states might correspond to building-up of motor program which is known to take place in premotor (PM) as well as supplementary motor area (SMA) [26]. Furthermore prefrontal cortex which has been assumed to play important roles in planning [27] might involve in searching optimal motor programs in terms of the initial state in PM or in SMA. It has been widely considered that PM involves more with generation of sensory-guided action while SMA does with the ones by internal drives [26]. Especially, the roles of mirror neurons [28] in ventral premotor (PMv) for action generation and recognition of the same action by others have drawn large attentions recently. Some emphasize mutual interactions between PMv and inferior parietal lobe (IPL) in generation of own actions and recognition of the same actions by other [29, 30]. They assume that PMv receives

integrated sensory information from IPL and PMv sends back the efference copy to IPL. Under these circumstances, the current paper focus on triangular functional relations among PFC, PMv and IPL in generation of own skilled actions and recognition of ones by others. One specific assumption in the current model is that IPL might play essential roles in acquiring the predictive model for well-practised actions by means of so-called the sensory-forward model[1]. More specifically, IPL might anticipate coming sensory flow in terms of proprioception and vision associated with a currently intended action. The assumption of the anticipation function for IPL might be unusual to some readers because the main role of IPL has been regarded as an integrator for different modalities of sensory perceptions [31]. However, there have been glowing evidences [32, 33, 34, 35] which could support the ideas of anticipatory functions in IPL.

The global picture of the proposed brain model would be like as shown in Figure 4. In this figure, PFC might set an intention for action program into

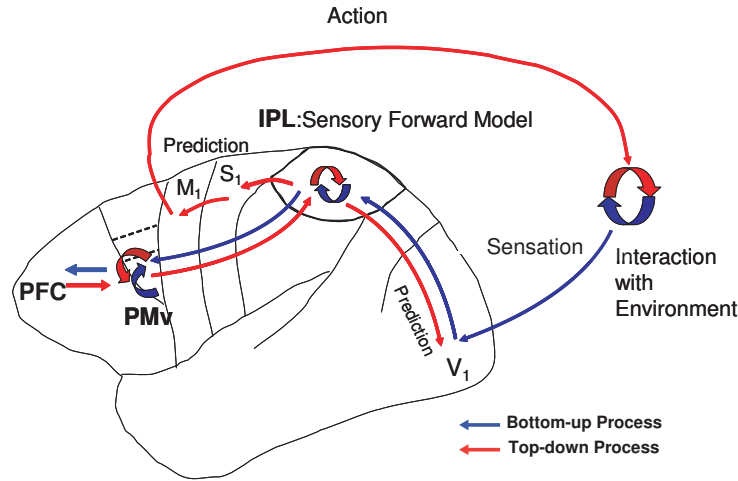


Figure 4: The proposed brain model where triangular functional relations among PFC, PMv and IPL are shown.

PMv by providing initial states. PMv might generate abstract sequencing of action primitives by means of its slow dynamics while IPL, which is bi-directionally connected to PMv, might generate detailed visuo-proprioceptive flow by anticipation by means of its fast dynamics. The prediction of proprioception in terms of body posture represented in the joint coordinate system

may provide target position at next moment to motor cortex or cerebellum in which necessary motor commands might be obtained by means of inverse computation. On the other hand, the prediction of visual image would generate visual mental imagery for future of which signal might descend the visual pathway in top-down manner. PFC might utilize the visual mental imagery generated in IPL for the purpose of searching the action programs in PMv which could provide the best match between the desired visual goal images and mentally simulated ones. The details of the brain model will be described in the later sections.

The current paper introduces a novel experiment concerning imitation of actions using a humanoid robot implemented with above mentioned neurodynamic model. We consider that there are two different ways of imitation. One is a trajectory level imitation, in which an imitator mimick an action trajectory of a demonstrator exactly[36]. This trajectory level imitation can be achieved by acquireing the ability of motion tracking, pose estimation, body correspondence and coordinate transformation. The other one is primitive level imitation, in which an imitator recognizes a demonstrated action as a sequence of action primitives, then the imitator maps demonstrator's primitives to own ones and generates own action primitive sequence. In the current paper, we are focusing on the latter case.

At first, the robot learns to bind observation of a human subject's (demonstrator's) actions with generation of own corresponding actions in supervised ways (see Figure5 for schematics). The learning is conducted for a set of different actions concerning manipulation of an object which can be decomposed into sequences of action primitives. The robot learning is conducted with having variations in relative observational angles from the robot to the human demonstrator as well as in object positions for the purpose of gaining generalization via accumulated learning. It is important to note that there could be a potential difference in visual perception during observation of demonstrator's action and generation of own corresponding action. So the robot should acquier more qualitative understanding of actions independent of the view, that is to find a set of action primitives shared by own generation case and observation of demonstrator's case as segmented from continuous perceptual trajectories even though their trajectory-level appearances are different between the two.

Then, an interesting test is to see how the robot can imitate novel actions shown by the demonstrator which are composed from prior-learned action primitives. This examination would address general questions in the imita-

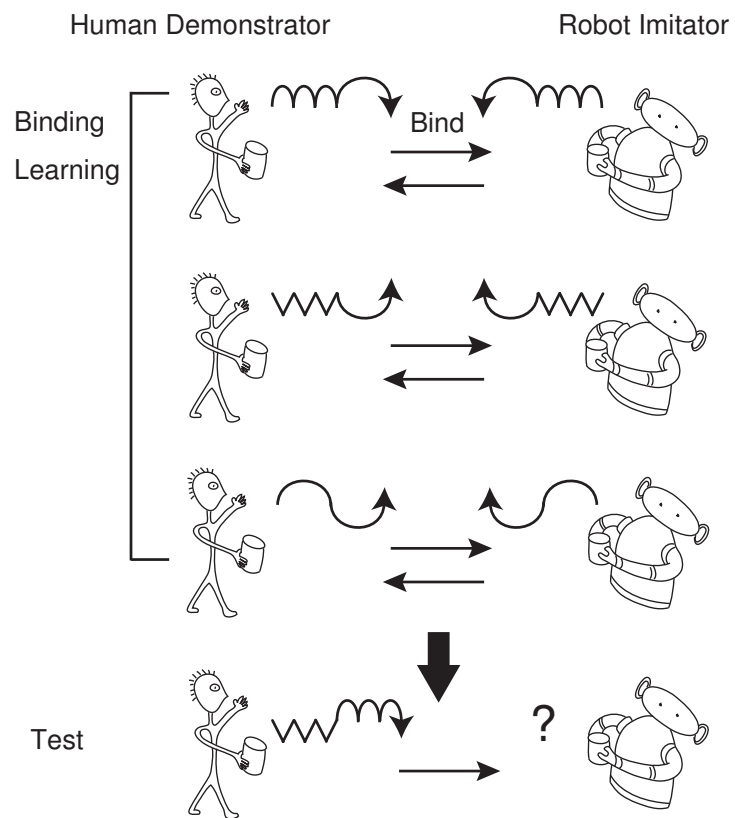


Figure 5: The imitation learning task.



tion learning community that how imitation can be conducted with qualitative understanding of other’s actions rather than merely mimicking physical action trajectories and how generalization via learning could be achieved [37, 38, 39]. The current paper attempts to show a possible account from our proposed neuro-dynamic systems perspective.

## 2. Model

In this section, our proposed neuro-dynamic model is explained in more details. Next section describes the over whole information flow by making possible correspondences to anatomical organization of brains. Especially, we describe two specific modes of on-line generation of anticipatory actions and off-line goal-directed plan generation mode. Then, more details of the neural network modelling scheme will be given in next-next subsection.

### 2.1. Brain model

#### 2.1.1. Generation of anticipatory actions

Figure 6 illustrates how anticipatory actions are generated by going through interactions between PMv and IPL. It is supposed that PMv receives inputs

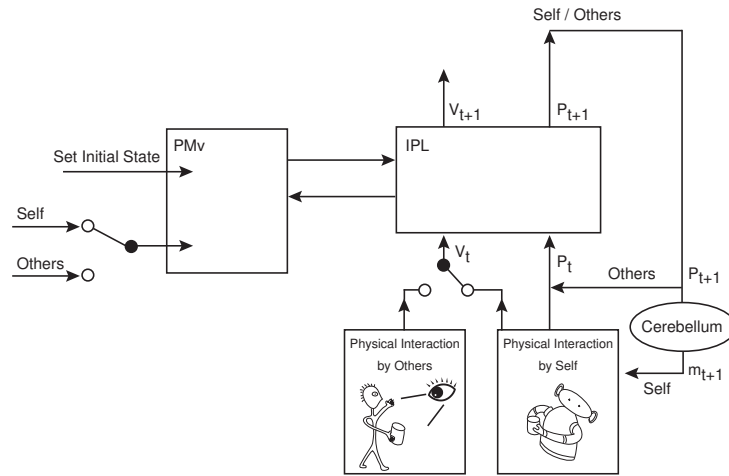


Figure 6: Generation of anticipatory behaviors via interactions between PMv and IPL.

from PFC which switch the system’s operational mode between the ‘self mode’ and the ‘demonstrator mode’. (Such neurons of distinguishing between own actions and demonstrator’s were found in parietal cortex [40] but

not yet in premotor.) The visuo-proprioceptive inputs accompanied by own actions are predicted in the 'self mode' while visual perception of observing demonstrator are predicted in the 'demonstrator mode'. PFC also sets the initial states (the initial potential values of some neuronal units) in PMv. After starting from the initial state, neuronal activation profiles change slowly in PMv which affects the forward dynamics in IPL by having mutual interactions between the two. In the 'self mode', coming visual inputs and proprioceptive inputs in terms of body posture are predicted in IPL based on the prior learning. The prediction of proprioception (posture in terms of joints angles) in next time step is utilized as the target to be achieved in next time step. Cerebellum or motor cortex might compute necessary motor torques to achieve the target for joint angles. On the other hand in the 'demonstrator mode', the visual inputs in next time step during observation of the demonstrator's action are predicted. Although there are prediction outputs for the proprioception also in the 'demonstrator mode', they are not utilized but are fed back as recurrent inputs.

### 2.1.2. Planning

In the current setting, planning is regarded as a process to determine the optimal action program which matches the distal visual images of mentally simulated with the goal image of specified (See Figure7). For this purpose,

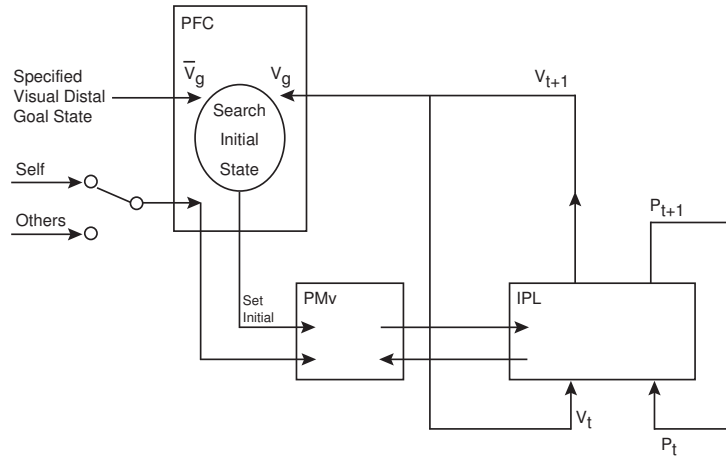


Figure 7: Planning by interactions among PFC, PMv and IPL.

the initial state of PMv is iteratively searched such that look-ahead prediction

of the visual image by IPL can fit with the goal image specified. The look-ahead prediction in IPL is conducted by means of mental imagery with closing the loops between the next step prediction and current step inputs for both vision and proprioception. The planning can be conducted both for own actions and demonstrator’s actions by switching between the ‘self’ and the ‘demonstrator’ modes.

## 2.2. Mathematical model

### 2.2.1. Overview

The multiple timescale recurrent neural network (MTRNN) [25] model used in this study is shown in Figure8. A small humanoid robot, which has

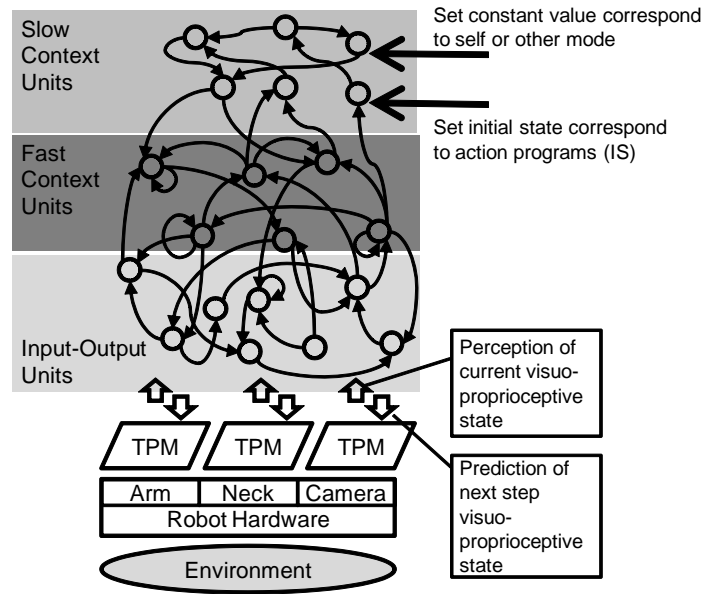


Figure 8: Neural network model interfacing with a robot

two 4-DOF arms, a 2-DOF neck and a stereo camera, is used in the role of a physical body interacting with actual environment. The neural network model receives the proprioceptive state ( $P_t$ ) representing the posture of arms in terms of the joint angles, the direction of the camera in terms of the neck joint angles ( $V_t^d$ ) and the visual image obtained with the camera ( $V_t^v$ ).

At first, current sensory inputs ( $P_t, V_t^d, V_t^v$ ) are pre-processed using topology preserving maps (TPM), which transforms a vector of continuous values

into neural population coding. After this transformation, input signals are sent to the MTRNN and the MTRNN predicts sensory inputs, which was encoded by neural population coding, for the next time step. This predicted signal is transformed inversely from neural population coding to continuous values  $(P_{t+1}, V_{t+1}^d, V_{t+1}^v)$  using TPM. Then the actual robot movement just follows this prediction of the proprioceptive state. In addition, the system can be operated also in so-called the closed-loop mode to generate “mental simulation” of sensory sequential patterns without generating physical movements by feeding back the prediction of the next sensory inputs to the current ones.

In our current model, the MTRNN consists of three parts, namely input-output units, fast context units and slow context units. Both the input-output units and fast context units correspond to the IPL. In addition, the slow context units corresponds to the PMv. The number of neural units is 100 for fast context, 50 for slow context respectively. Among the slow context units, there are four special neuronal units. Two of them are used to specify action programs. The initial states of these neurons (IS) take the same value both in generating a particular action in the ‘self mode’ and in generating “mental imagery” of observing the corresponding action by the demonstrator. Each initial state for corresponding action sequence is self-determined in the learning process as bound to be the same for the both modes. Other two neurons function as a switch between ‘self mode’ and ‘demonstrator mode’. The neuronal activation of these switching units are set different value, which takes constant value from the start to the end of the action sequence, by PFC depending on the operational modes between the ‘self mode’ and the ‘demonstrator mode’.

In our current model, PFC have two functions. The first one is to remember initial states of the PMv, which were self-determined in the learning process, as a database. When the robot is required to generate action sequence, which the robot have already experienced, PFC recall the initial state from the memory and set the initial state of PMv. The other is to find the initial state of PMv corresponding to the optimal action program to achieve the given goal state. This search algorithm is described in the following section.

### *2.2.2. Sparse encoding of input-output*

Input to the neural network are sparsely encoded in the form of a population coding using topology preserving maps (TPM). The current model

has three TPMs, one map corresponding to proprioceptive state  $P$ , one map corresponding to neck joint angle  $V^d$  and one map corresponding to camera image  $V^v$ . The TPM is a type of a neural network that produces a discretized representation of the input space of training samples. This sparse encoding of input trajectories reduces the overlaps of input sequence, so that it improves learning ability of the MTRNN as described in [25]. The size of TPMs are 100( $10 \times 10$ ) for proprioception, 36( $6 \times 6$ ) for neck joint angle and 144( $12 \times 12$ ) for camera image respectively.

In the current study, TPMs are trained in advance of MTRNN training using conventional unsupervised learning algorithm[41]. Training data set of TPMs included all teaching signals for the MTRNN. The TPM transformation is described by following fomula,

$$q_{i,t} = \frac{\exp\left\{-\frac{\|k_i - k^{teach}\|^2}{\sigma}\right\}}{\sum_{j \in Z} \exp\left\{-\frac{\|k_j - k^{teach}\|^2}{\sigma}\right\}} \quad (1)$$

where if  $i \in P$ , then  $Z = P$  and  $k^{teach} = P_t$ , if  $i \in V^d$ , then  $Z = V^d$  and  $k^{teach} = V_t^d$ , if  $i \in V^v$ , then  $Z = V^v$  and  $k^{teach} = V_t^v$ .  $\sigma$  is a constant value, which indicating the shape of the distribution of  $q_{i,t}$ , set at 0.1 in the current study.  $k_i$  is reference vector of  $i$ -th neural unit in a TPM.

The MTRNN generates prediction of next step sensory visuo-proprioceptive state based on the aquired forward dyanmics described later. This prediction generated by MTRNN can be assumed to correspond to an activation probability distribution over the TPMs. It is inversely transformed to a visuo-proprioceptive input signal using the same TPMs:

$$k^{out} = \sum_{i \in Z} y_{i,t} k_i \quad (2)$$

where if  $i \in P$ , then  $Z = P$  and  $k^{out} = P_{t+1}$ , if  $i \in V^d$ , then  $Z = V^d$  and  $k^{out} = V_{t+1}^d$ , and if  $i \in V^v$ , then  $Z = V^v$  and  $k^{out} = V_{t+1}^v$ .

### 2.2.3. Forward generation of MTRNN

The MTRNN [25] is a type of the continuous time recurrent neural network (CTRNN). In the CTRNN, neuronal activations are calculated using conventional firing rate model shown in Eq.3.

$$\tau_i \frac{du_{i,t}}{dt} = -u_{i,t} + \sum_j w_{ij} x_{j,t} \quad (3)$$

$$y_{i,t} = \begin{cases} \frac{\exp(u_{i,t})}{\sum_{j \in Z} \exp(u_{j,t})}, & \text{if } i \in Z \\ \frac{1}{1 + \exp(-u_{i,t})}, & \text{otherwise} \end{cases} \quad (4)$$

$$x_{i,t+1} = \begin{cases} q_{i,t+1}, & \text{if } i \in Z \\ y_{i,t}, & \text{otherwise} \end{cases} \quad (5)$$

where  $u_{i,t}$  is the potential of each  $i$ -th neural unit at time step  $t$ ,  $x_{j,t}$  is the neural state of the  $j$ -th unit and  $w_{ij}$  is synaptic weight from the  $j$ -th unit to the  $i$ -th unit and  $Z$  is  $P$  or  $V^d$  or  $V^v$  and  $q_{i,t}$  is input vectors using TPMs.

The characteristic of the MTRNN is that they consist of neurons of multiple time-scales dynamics. The MTRNN, which is used in the current study, consisted of input-output units and non-input-output units, the latter referred to as context units. Context units are divided into two groups based on the value of time constant  $\tau$ . The first group consisted of fast context units with small time constant ( $\tau = 2$ ), whereas the second group consisted of slow context unit with a large time constant ( $\tau = 10$ ). This characteristic enables the MTRNN to articulate complex visuo-proprioceptive sequences into reusable primitive patterns by self-organizing a functional hierarchy of them while learning. The effect of having multiple time-scale dynamics among groups is roughly summarized as follows: neurons with fast time constant assemble a set of primitives and neurons with slow time constant "manipulate" them in various sequences depending on IS.

IS values are set by PFC before forward generation. Those values are self-determined under the condition that they take the same value when generating the particular action in the 'self mode' and the corresponding mental imagery in 'demonstrator mode' through the learning process. The potential values  $u$  of neural units corresponding to self-other switch are set 100 in 'self mode', whereas the values are set -100 in 'demonstrator mode'. During the learning phase of the MTRNN, the same potential values are used.

#### 2.2.4. Learning of own and demonstrator's action sequences

In our current study, the TPMs are trained in advance of the MTRNN training using conventional unsupervised learning algorithm. The goal of training the MTRNN is to find the optimal values of the connective weights and IS which minimize the value of E, which is defined as the learning error. We define the learning error using Kullback-Leibler divergence as a following

equation,

$$E = \sum_t \sum_{i \in IO} y_{i,t}^* \log\left(\frac{y_{i,t}^*}{y_{i,t}}\right) \quad (6)$$

where  $IO$  represents input-output units,  $y_{i,t}^*$  is the desired activation value of the output neuron at time  $t$  and  $y_{i,t}$  is the activation value of the output neuron with current connective weights and IS. As training method, we use the general Back Propagation Through Time (BPTT) algorithm [42]. Using the BPTT algorithm, the network can reach their optimal levels for all given teaching sequences by updating the connective weights in the opposite direction of the gradient  $\partial E/\partial w$ .

$$w_{ij}(n+1) = w_{ij}(n) - \alpha \frac{\partial E}{\partial w} \quad (7)$$

where  $\alpha$  is the learning rate constant, and  $n$  is an index representing the iteration epoch in the learning process. Note that there are both own action sequence and demonstrator's action sequence in the training data set. However, there are no teaching signals for proprioceptive sequence when the neural network is trained with demonstrator's action sequence, and so the error is calculated only with the direction of the camera ( $V_t^d$ ) and visual image ( $V_t^v$ ) for demonstrator's action sequences.  $\frac{\partial E}{\partial w}$  is given by:

$$\frac{\partial E}{\partial w} = \sum_t \frac{1}{\tau_i} \frac{\partial E}{\partial u_{i,t}} y_{j,t-1} \quad (8)$$

and  $\frac{\partial E}{\partial u_{i,t}}$  is recursively calculated from the following recurrence formula

$$\frac{\partial E}{\partial u_{i,t}} = \begin{cases} y_{i,t} - y_{i,t}^* + (1 - \frac{1}{\tau_i}) \frac{\partial E}{\partial u_{i,t+1}} & \text{if } i \in IO \\ \sum_{k \in N} \frac{\partial E}{\partial u_{i,t+1}} \left[ \delta_{ik} (1 - \frac{1}{\tau_i}) + \frac{1}{\tau_k} w_{ki} f'(u_{i,t}) \right] & \text{if } i \notin IO \end{cases} \quad (9)$$

where  $f'()$  is the derivative of the sigmoidal function and  $\delta_{ik}$  is Kronecker delta.

In addition, an update of the initial context unit values corresponding to action sequence (IS) is obtained simultaneously by utilizing the delta error  $\frac{\partial E}{\partial u_{i,0}}$  back-propagated through time steps to the context unit at the initial step.

$$c_{i,0}(n+1) = c_{i,0}(n) - \alpha \frac{\partial E}{\partial u_{i,0}} \quad (10)$$

where  $c_{i,0}$  is the potential of  $i$ -th neural unit. As described in previous section, the initial states correspond to same action programs should be the same value either in the 'self mode' or 'demonstrator mode' in the proposed "bind" learning schema. To satisfy this condition, the initial state value correspond to same action programs are averaged among the training data set after applying the Eq 10. It is summarized that the learning proceeds with dense interaction between the top-down and the bottom-up pathways where "imagery" of visuo-proprioceptive (VP) sequences generated in the top-down pathway by means of current weights and IS values are iteratively encountered by the actual VP sequences of the bottom-up. As the generated error is propagated through all neurons including the fast and slow parts, the synaptic weights in the whole network can be updated.

### 2.2.5. Goal-directed planning

When the goal state is given by a experimenter, in terms of distal visuo-proprioceptive state  $(P_{goal}, V_{goal}^d, V_{goal}^v)$ , IS value to generate VP sequences that include the goal state in their distal steps, is searched. The search is conducted by means of BPTT algorithm [42] but without modulating the synaptic weights. Firstly, the MTRNN generates "imagery" of VP sequence by closed-loop operation with temporarily set IS values. At this time, neuronal activation of context neurons corresponding to self-other switch are set to 'self mode'. Then, the Euclidean distance between the given goal state and the state generated in the imagery VP sequences are calculated for all time steps in the sequences, and the particular time step where the calculated Euclidean distance becomes minimum are marked up. The error is back-propagated from this marked step to the initial step in the sequence and the IS values are updated using Eq 10. The search is started by setting the initial IS values randomly and it is iterated until the error is minimized. Then the plan is enacted by the robot in the physical environment with activating the network forward dynamics with setting IS with the obtained values in the search.

## 3. Experiment

### 3.1. Task design

The experiment was carried out by using a small humanoid robot named HOAP3. The robot was fixed to a chair, and a workbench was set up in front of the robot. A movable rectangular solid object which is painted



half blue and half red was placed on the workbench. The neck joints are controlled by PID controller, which is programmed to track a red-color object to be centered in the retinal image. Inputs to the neural model are angles of arm joints  $P_t$  (8 dimensional vector) and neck joints  $V_t^d$  (2 dimensional vector) and the visual perception  $V_t^v$  ( $16 \times 12$  dimensional vector). Each pixel component of the visual perception vector can take one of four possible values (0, 0.375, 0.625, 0.875) depending on the color (others, green, blue, red) of corresponding retinal cell, respectively.

The compositional actions used in current experiment is shown in Figure 9. As shown in this figure, compositional actions assumed in the experimental task had a temporal structure which could be described by a path of state transitions with branching. The first half of the action is moving the object to the right (R) or moving the object to upward (U), whereas the last half of the action is knocking over the standing object to lie by left arm (K) or moving the object to the left (L). Totally, as combinations of the first half and the last half, there are four actions, i.e (RK), (RL), (UK), (UL).

In order to obtain a teaching signal of self actions, the human tutor guided both hands of the robot along the trajectory of the action without providing explicit cues of showing segmenting points for the primitives. In these actions, the robot started to move from the same home position so that the MTRNN can not utilize the initial posture difference information to generate difference action programs. As the robot hands were guided along the trajectory, the joint angles at each step were recorded and their sequences were used as teaching sequences. For each actions, the robot was tutored 3 times with changing the initial object position as 2cm left from the original position, the original one and 2cm right from original one. Meanwhile, other’s action sequences were obtained by showing the corresponding demonstrator’s actions to the robot. The sensory information during this observation of the demonstrator’s actions consists of the vision inputs and the head direction of the robot. For each action by the demonstrator, the action sequence was recorded 3 times with changing the relative position between the demonstrator and the robot as located in front, in 15 degree right and in 15 degree left. In the current setup of the experiment, the demonstrator demonstrated action sequences using only one hand. Therefore the appearances of the same actions by the demonstrator and the robot could be significantly different as shown in Figure 10.

At first, we examine how the robot can imitate novel actions shown by the demonstrator by utilizing prior-learned behavior skills. For this purpose the

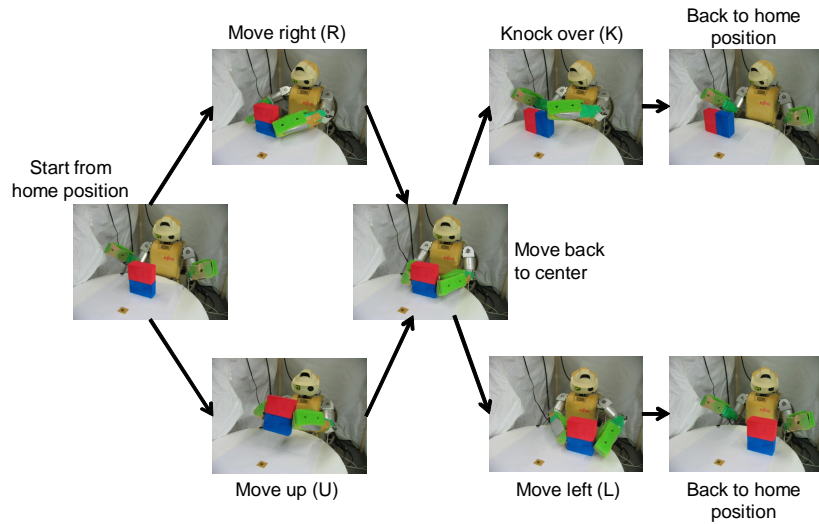


Figure 9: Task design: Robot was trained for three task trajectories of (R)(K), (U)(K) and (U)(L) with starting and going back to home position. Then, it was tested if the robot can generate (R)(L) task trajectory by observing its corresponding action by demonstrator.

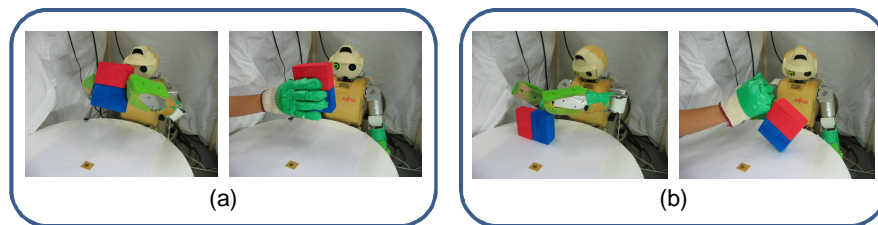


Figure 10: These pictures show the appearance of compositional actions performed by the robot or human imitator. In (a) the robot move the object upward using both hands, meanwhile the demonstrator performed the corresponding action using only one hand. In (b) the robot or demonstrator knock over the object from its left-hand side to its right-hand side.

robot was trained for three actions including (RK), (UK) and (UL) by means of the bind training of the observation of the demonstrator’s actions and the own ones. After the training, another observation of the demonstrator doing the action sequence (RL) was added to the training data set and the corresponding initial state values was calculated. In the test generation, it is examined if the robot can successfully generate the novel own action sequence of (RL) by using the initial state obtained by learning of the observation of its corresponding action by the demonstrator.

In the next, we examine how plans for actions can be generated in the proposed scheme. The test was conducted using the network trained in the previous experiment. In the planning, the action program should be made to achieve the goal state of the object being located in the left hand side from the initial state of the object standing on the workbench with the robot’s posture in its home position. Note that this goal state can be achieved by generating either action sequence of (RL) or (UL). In this experiment, action plans were generated by searching for the initial state from which the sensory-forward model can bring the best match between the specified goal state and the predicted one in the distal step as described in the previous section.

### 3.2. Results

#### 3.2.1. Regeneration of the trained actions

All three TPMs were pre-adapted before the training of the MTRNN utilizing the sensory patterns including self and other’s action sequences. For the training of MTRNN, BPTT was iterated for 12000 epochs with randomly set initial weights, and it took about 5 hours in the current setting. The mean square error (MSE), which is the average square error per input-output neurons per step over all teaching sequences, conversed to 0.000733.

After the training, the MTRNN was tested to regenerate trained actions by the robot by setting the IS with the corresponding values determined through the learning process. Each action was tested with varying the object position among three of originally trained positions and two of novel positions which were 1 cm left and right from the center. Each trial was repeated for three times. The result of the experiment show that the robot could regenerate trained actions with a success rate of 100%. Figure11 shows the time development of the actions in which the object is placed in the original position. The upper graph shows the encoder readings of four joint angles of the right arm which are shown as unified in the range between 0.0 and 1.0. The lower graphs show the activation history of the IS neurons by

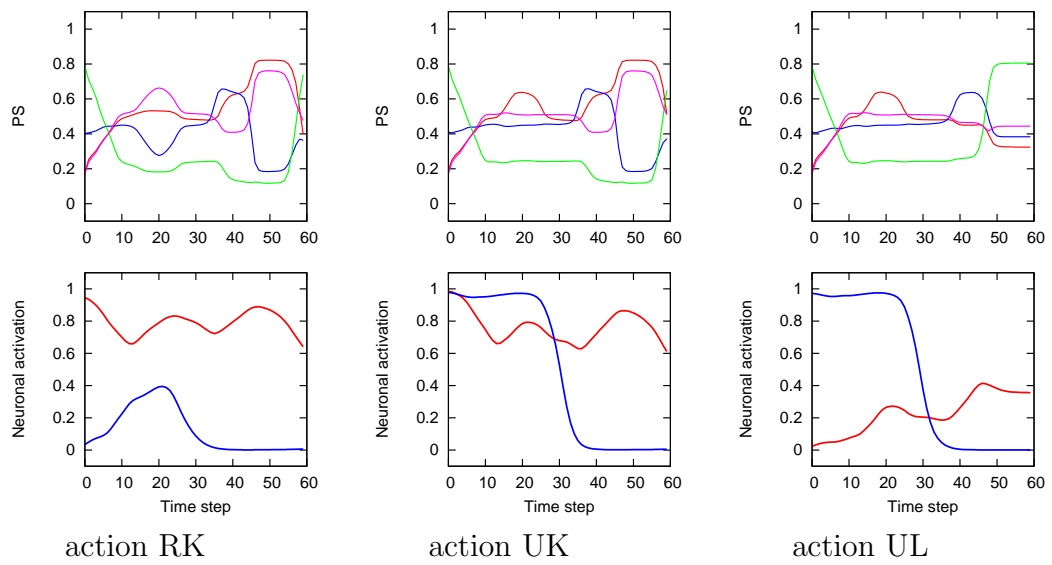


Figure 11: Time development of three actions generated. Each pair of three graphs corresponding to the three actions as described in Figure9. The upper graphs show the proprioceptive state (PS) for the four joint angles in the right arm. The lower graphs show the activation history of the IS neurons.

red and blue lines whose initial values were self-determined in the learning. As shown in this figure, the initial values of the IS neurons are different depending on the action sequence. In order to analyze the relationship between the IS and corresponding action sequences generated, we constructed the initial state map which shows the mapping from two-dimensional initial state to the resulting imagery VP sequences. The result can be seen in Figure 12. We compared a mental image, which is generated with each initial state in the grid, with trained action primitive pattern and classified the mental image manually. Each grid shows a label of representing generated combination of action primitive pattern, where (R), (L), (K), (U) denoting classified primitives, (X) denotes unclassifiable “distorted” ones. It is seen that three of trained combinations of the primitives, namely (UL), (UK) and (RK) appear as large clusters while distorted one appear in small regions.

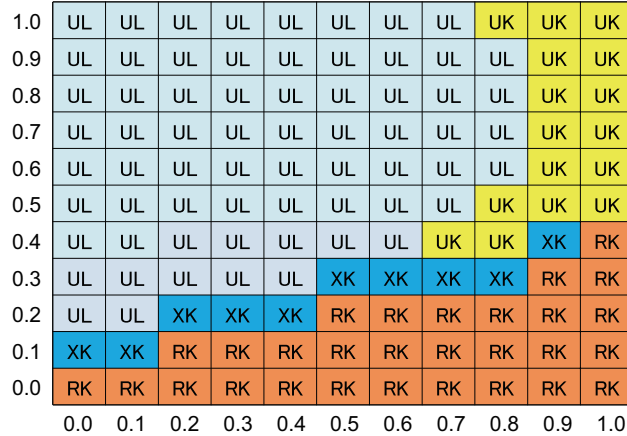


Figure 12: The initial state map obtained after the first training of the MTRNN. Each grid is labeled in accordance to the generated combination of action primitive pattern, where (R), (L), (K), (U) denoting the classified primitives, (X) denotes unclassifiable distorted one

### 3.2.2. Generation of novel action by imitation

After the first training, observation of the demonstrator doing the novel action sequence (RL) was added to the training dataset. Then BPTT learning was iterated for 3000 epochs on the trained synaptic weights which had been obtained through the first training, and it took about 75 minutes. As a result the MSE conversed to 0.000591 with the new dataset including the

observation of novel action sequence. The imitational action sequence was generated by using the initial state determined for the novel action sequence with the 'self mode'. The generated action sequence was tested with varying the object position that was the same condition of regeneration of trained action sequences. The result of this experiment shows that the robot could generate imitational action sequences with a success rate of 100%. The time development of the imitational action sequence in which the object is placed in the original position is shown in Figure 13.

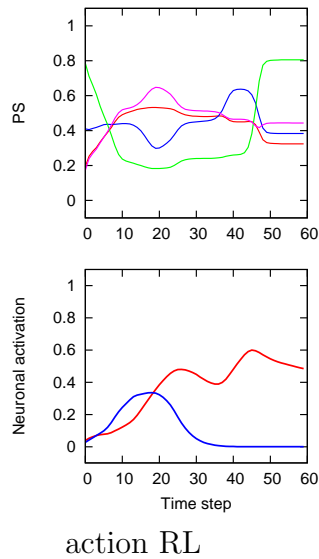


Figure 13: Time development of the imitational action. The upper graph shows the proprioceptive state (PS) for the four joint angles in the right arm. The lower graph shows the activation history of the IS neurons.

When looking at Figure11 and Figure13, we can find that the IS values shown by the initial value of red and blue lines take different value depending on the combination of action primitives. When the first half of the action sequence is (U) the blue line in the lower graph starts from lower value, whereas it starts from higher value when the first half of the action sequence is (R). Likewise the red line in the lower graph starts from lower value when the last half of the action sequence is (L) and it becomes (K) when the red line starts from higher value. This observation suggests that compositional action primitive sequences are successfully embedded in the slow dynamics

of the IS neurons.

Figure 14 indicates the initial state map obtained after the additional training. With comparing with Figure 12, the novel cluster corresponding to the action sequence (RL) appears in the left bottom of this map while positions of the clusters corresponding to previously learned action sequences (UL), (UT), (RT) change only slightly. This result is interpreted that a novel own action program that corresponds to observation of demonstrator’s action can be generated by utilizing prior learned structures of action primitives.

1.0	UL	UL	UL	UL	UL	UL	UL	UL	UL	UK	UK
0.9	UL	UL	UL	UL	UL	UL	UL	UL	UL	UK	UK
0.8	UL	UL	UL	UL	UL	UL	UL	UL	UL	UK	UK
0.7	UL	UL	UL	UL	UL	UL	UL	UL	UL	UK	UK
0.6	UL	UL	UL	UL	UL	UL	UL	UL	UL	UK	UK
0.5	UL	UL	UL	UL	UL	UL	UL	UL	UL	UK	UK
0.4	UL	UL	XK	XK	XK	XK	XK	XK	XK	XK	RK
0.3	RL	RL	RL	RL	RL	RK	RK	RK	RK	RK	RK
0.2	RL	RL	RL	RL	RL	RK	RK	RK	RK	RK	RK
0.1	RL	RL	RL	RL	RL	RK	RK	RK	RK	RK	RK
0.0	RL	RL	RL	RL	RL	RL	RK	RK	RK	RK	RK
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

Figure 14: Initial state map generated after the additional training of novel action sequence (RL)

### 3.2.3. Planning of action for given goal state

The action plans to achieve the specified goal state as described in the previous section were generated by searching the optimal IS values. The search calculation was iterated 2000 epochs with the learning rate  $\alpha = 0.2$ , and it took about 50 minutes in the current setting. As had been expected, two distinct action plans of (RL) and (UL) were generated in terms of imagery VP sequences depending on the initial setting of the IS values in the search process. Those two typical VP sequences are shown in Figure 15 for RL and Figure 16 for UL. In these figures, imagery sequences generated in both 'self mode' and 'other's mode' are shown. These figures indicate that two possible imagery for own action as well as for other's in achieving the specified goal are successfully generated by means of planning. It was also confirmed that

the robot can successfully execute those plans in its physical behaviors by setting the IS with those values obtained in the search.

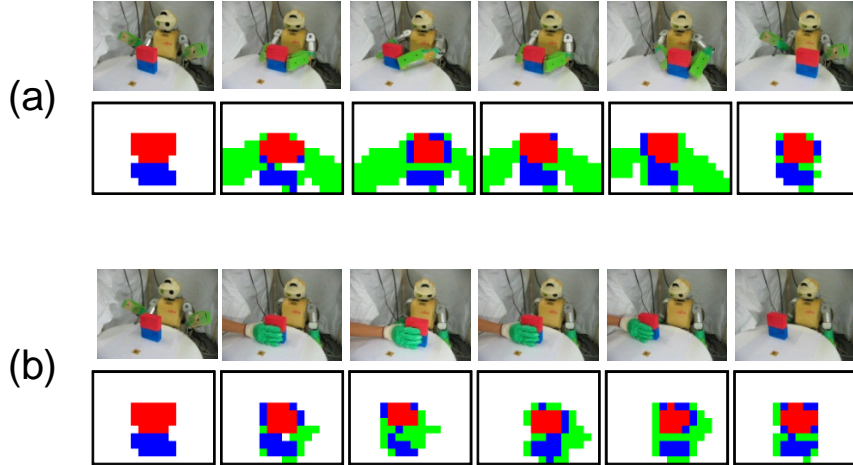


Figure 15: The visual imagery corresponding to RL generated as a result of planning. In (a) visual imagery for own action in the lower row and actual execution of the plan in the upper row. In (b) visual imagery for demonstrator's action in the lower row and actual demonstrator's action demonstrated to the robot in the upper column.

To carry out more detailed analysis, we constructed a map which shows the mapping from the initial IS (IIS) values, which is set to the IS in the onset of the search, to action sequences generated from the converged IS value in the search. The IIS values are adjusted in the range between 0.0 and 1.0 with 0.1 interval. The result is shown in Figure 17. It can be seen that the search process can generate correct action plans of either (RL) or (UL) with setting the IIS at most points in the space. Also, we can see two clusters of (UL) in the upper part and of (RL) in the lower part in the space those overlap largely with the corresponding clusters shown in Figure 14. This analysis indicates that goal-directed planning can be performed quite robustly.

#### 4. Discussion

The current paper presented a neuro-dynamic systems approach to the problem of generating/understanding actions of own/demonstrator's by introducing our synthetic robot modeling study. In the following, we show our



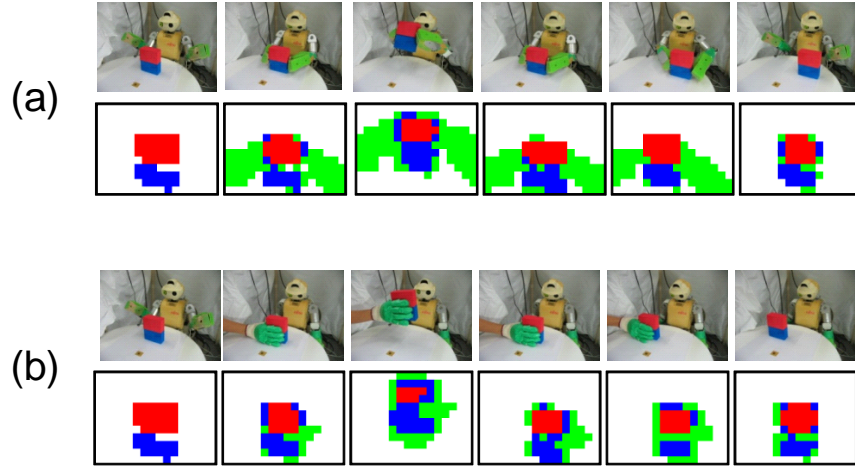


Figure 16: The visual imagery corresponding to UL generated as a result of planning. See the caption of Figure 14.

1.0	UL	UL	UL	UL	UL	UL	UL	UL	UL	UL	R
0.9	UL	UL	UL	UL	UL	UL	UL	UL	UL	U	R
0.8	UL	UL	UL	UL	UL	UL	UL	UL	UL	RL	RL
0.7	UL	UL	UL	UL	UL	UL	UL	UL	UL	RL	R
0.6	UL	UL	UL	UL	UL	RL	UL	UL	UL	UL	RL
0.5	UL	RL	RL	RL	RL	UL	UL	UL	RL	RL	RL
0.4	RL	RL	RL	UL	RL	RL	RL	UL	RL	RL	R
0.3	RL	RL	RL	UL	UL	RL	RL	RL	RL	RL	R
0.2	RL	RL	RL	RL	RL	RL	RL	RL	RL	RL	R
0.1	RL	RL	RL	RL	R	RL	RL	RL	RL	R	RL
0.0	RL	RL	RL	RL	RL	R	R	RL	RL	RL	R
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

Figure 17: Initial IS map. This map shows the mapping from the initial IS (IIS) value to imagery sequences generated with the converged IS value in the search. Each grid indicates a label of imagery sequences.

interpretation and discussion about the obtained experimental results qualitatively with focusing on the problems of generalization in imitation learning and creations of novel actions. We will also discuss future related studies.

#### *4.1. Generalization in imitation learning*

It is considered that imitation learning require at the least two classes of generalization [37, 38, 39]. One is generalization over sensory-motor trajectory variances such as caused by differences in object size, position and shape to be manipulated or by view differences in observing “demonstrator”. The current study showed that such trajectory level variances related to object positions and views on “demonstrator” can be resolved by introducing unified intentional states in terms of the initial internal states values over such variances by means of the proposed binding learning scheme. Such intentional states might be hosted by mirror neurons [28] in PMv.

The other class of generalization is on cognitive level which concerns on more qualitative understanding of actions by own and demonstrator’s which do not necessarily depend on exact action trajectories. Understanding of actions require understanding of multiple possible ways of achieving the same goal with abstraction. This requires decomposition of original visuo-proprioceptive trajectories into reusable primitives and their recombinations into multiple pathways to achieve the same goal. Our experimental results showed that such abstraction can be realized by learning in multiple time-scales neuronal dynamics and that the action planning scheme of searching the initial state can construct mental imagery of multiple pathways to achieving the same goal.

On this aspect, some may claim that the same function could be accomplished by some hybrid system approaches [7, 8] of interfacing symbol level constructed by some machine learning scheme and analog representation in sensory-motor level by using neural nets of fuzzy systems. We, however, argue that such approaches may suffer from the symbol grounding problems [9] because the interface between those level can be just arbitrary. On the contrary, in our proposed dynamical systems approach, there are no explicit separation of symbol level and sensory-motor level because the whole cognitive-behavioral function is realized by one body of neuro-dynamical system. When seamless interactions are iterated between the top-down goal-directed intentionality and the bottom-up sensation of the reality in the adopted neuro-dynamical system through repeated learning, cognitive constructs of being

compositional, yet situated to sensory-motor reality should emerge upon the self-organized dynamical structures.

#### *4.2. Creation of novel actions*

One of interesting question to ask might be how diversity of novel action programs can be generated by utilizing and modifying prior-learned action scheme. Our experiment showed that after the robot observes demonstrator’s action, its corresponding own action, which is novel for the robot, can be successfully generated by adequately mapping the observed one to the own one by utilizing the prior-learned own primitive. The initial state mapping analysis shown in Figure 14 indicated that relatively large clusters appear for the IS of the novel combination of action primitives as well as for that of other trained combinations but that only small region does for that of distorted combination of primitives. It was, however, shown in our preliminary experiment that such novel combination cannot be generated anymore unless the observation of corresponding demonstrator’s action is binded in the 2nd time learning. On the contrary, our prior study [24] using MTRNN indicated that diverse novel combinations of action primitives including that of distorted ones can be generated by introduction of chaotic dynamics in the slow dynamics part in the network. These results suggest that diversity of creating novel actions of combining prior-learned primitives can be drastically increased by introduction of chaos. It should be, however, also true that introduction of chaos may lead to unstable situations where useless distorted actions are more likely to be generated. It is speculated that “creativity” for generating novel actions could face the dilemma between flexibility and stability.

#### *4.3. Future studies*

Future studies should extend the vision system for the purpose of scaling the task complexity. In the current experiments, variance of relative position between the robot and the demonstrator was limited in narrow range of plus-minus 15 degrees. Therefore, the visual sensation of the demonstrator can be easily mapped to that of the imitator by nonlinear function of adopted network. It is, however, not certain that the same scheme can be applied to the case of larger range of view variances. In such case, vision processing known as “mental rotation” [43] might be required.

In the course of scaling task complexity, the mechanisms of visual attention and saccading might be required to deal with manipulation of more than

one object. In our current setting, because it was assumed that there is only a single object to be manipulated, no attention switching mechanism was required. It is left for future studies to investigate how the scheme of attention switching accompanied by saccading can be attained through learning.

Another line of interesting future investigations might be comparison between the stochastic learning approach using Bayesian formulation [44] and deterministic dynamical systems approach. Recently, fast progress in the stochastic computational models such as hidden Markov model (HMM) attract large attention in many research area. The action imitation task had been also achieved by using the HMM[45]. One interesting recent results by Namikawa and Tani [46] is that forward model by CTRNN can learn probability of branching of primitives as embedded in its self-organized chaos. Such capability of mimicking stochastic sequences by deterministic dynamic neural network model should be qualitatively examined with comparison with various stochastic modelling scheme.

## 5. Conclusion

The current paper presented a synthetic robot modeling study which accounts for how actions by others can be recognized and imitated with compositionality in terms of the mirror neuron system by utilizing a dynamic neural network model of MTRNN. The method was evaluated in the experiments using a small humanoid robot implemented with the model. The results showed that (1) all of tutored actions can be regenerated robustly by generalizing the variance of the position of the target object, that (2) novel actions demonstrated by the demonstrator can be imitated by means of combining prior acquired primitives into the corresponding actions of self-generated, and that (3) when the goal state is given in terms of own visuo-proprioceptive state, multiple action plans which achieve the given goal state can be found robustly through iterative search. It is finally concluded that, by having seamless and dense interactions between the top-down goal-directed intentionality and the bottom-up perception of the reality through repeated learning of own experiences, cognitive constructs of being compositional, yet situated to sensory-motor reality can emerge as the results of self-organization of adequate neuro-dynamical structures.

## Acknowledgement

This work was supported in part by funds from the Grant-in-Aid for Scientific Research in Priority Areas “Emergence of Adaptive Motor Function through Interaction among the Body, Brain and Environment: A Constructive Approach to the Undergoing of Mobiligence” from Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

## References

- [1] J. Tani, R. Nishimoto, R. Paine, Achieving ”organic compositionality” through self-organization: Reviews on brain-inspired robotics experiments, *Neural Networks* 21 (2008) 584–603.
- [2] G. Evans, Semantic theory and tacit knowledge, in: S. Holzman, C. Leich (Eds.), *Wittgenstein: To Follow a Rule*, Routledge and Kegan Paul, London, 1981, pp. 118–137.
- [3] M. Arbib, Perceptual structures and distributed motor control, in: *Handbook of Physiology: The Nervous System, II. Motor Control*, MIT Press, Cambridge, MA, 1981, pp. 1448–1480.
- [4] M. Mataric, Integration of representation into goal-driven behavior-based robot, *IEEE Trans. Robotics and Automation* 8 (3) (1992) 304–312.
- [5] S. Harnad, Grounding symbolic capacity in robotic capacity, in: L. Steels, R. Brooks (Eds.), *The “artificial life” route to “artificial intelligence.” Building situated embodied agents*, Lawrence Erlbaum, New Haven, 1993.
- [6] A. Cangelosi, A. Greco, S. Harnad, From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories, *Connection Science* 12 (2) (2000) 143–162.
- [7] R. Sun, *Connectionist implementationalism and hybrid systems*, Nature Publishing Group MacMillan, 2002.
- [8] P. Doherty, The witas integrated software system architecture, *Linköping Electronic Articles in Computer and Information Science* 4 (17).

- [9] S. Harnad, The symbol grounding problem, *Physica D* 42 (1990) 335–346.
- [10] J. Tani, Model-based learning for mobile robot navigation from the dynamical systems perspective, *IEEE Trans. on SMC (B)* 26 (3) (1996) 421–436.
- [11] J. Tani, An interpretation of the "self" from the dynamical systems perspective: a constructivist approach, *Journal of Consciousness Studies* 5 (5-6) (1998) 516–42.
- [12] J. Tani, Learning to generate articulated behavior through the bottom-up and the top-down interaction process, *Neural Networks* 16 (2003) 11–23.
- [13] B. I. Hao, *Elementary Symbolic Dynamic and Chaos in Dissipative System*, World Scientific, Singapore, 1989.
- [14] J. Crutchfield, Inferring statistical complexity, *Phys Rev Lett* 63 (1989) 105–108.
- [15] J. Elman, Finding structure in time, *Cognitive Science* 14 (1990) 179–211.
- [16] J. Pollack, The induction of dynamical recognizers, *Machine Learning* 7 (1991) 227–252.
- [17] J. Tani, N. Fukumura, Embedding a grammatical description in deterministic chaos: an experiment in recurrent neural learning, *Biological Cybernetics* 72 (1995) 365–370.
- [18] K. Doya, S. Yoshizawa, Memorizing oscillatory patterns in the analog neuron network, in: *Proc. of 1989 int. joint conf. on neural networks*, Vol. 1, Washington, D.C., 1989, pp. 27–32.
- [19] R. J. Williams, D. Zipser, A learning algorithm for continually running fully recurrent neural networks, *Neural Computation* 1 (2) (1989) 270–280.
- [20] D. M. Wolpert, M. Kawato, Multiple paired forward and inverse models for motor control, *Neural Networks* 11 (7-8) (1998) 1317–1329.

- [21] J. Tani, S. Nolfi, Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems, *Neural Networks* 12 (1999) 1131–1141.
- [22] J. Tani, M. Ito, Y. Sugita, Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using rnnpb, *Neural Networks* 17 (2004) 1273–1289.
- [23] R. Nishimoto, J. Namikawa, J. Tani, Learning multiple goal-directed actions through self-organization of a dynamic neural network model: A humanoid robot experiment, *Adaptive Behavior* 16 (2-3) (2008) 166–181.
- [24] H. Arie, T. Endo, T. Arakaki, S. Sugano, J. Tani, Creating novel goal-directed actions at criticality: A neuro-robotic experiment, *New Mathematics and Natural Computation*.
- [25] Y. Yamashita, J. Tani, Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment, *PLoS Computational Biology* 4 (11).
- [26] M. I. H. Mushiake, J. Tanji, Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements, *Journal of Neurophysiol* 66 (3) (1991) 705–718.
- [27] J. Fuster, *The Prefrontal Cortex*, Raven Press, New York, 1989.
- [28] G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, Premotor cortex and the recognition of motor actions, *Cognitive Brain Research* 3 (2) (1996) 131–141.
- [29] A. H. Fagg, M. A. Arbib, Modeling parietal-premotor interactions in primate control of grasping, *Neural Networks* 11 (1998) 1277–1303.
- [30] E. Oztop, D. Wolpert, M. Kawato, Mental state inference using visual control parameters, *Cognitive Brain Research* 22 (2005) 129–151.
- [31] C. Colby, J. Duhamel, M. Goldberg, Ventral intraparietal area of the macaque: anatomic location and visual response properties, *Journal of Neurophysiology* 69 (1993) 902–914.

- [32] E. Eskandar, J. Assad, Dissociation of visual, motor and predictive signals in parietal cortex during visual guidance, *Nature Neuroscience* 2 (1999) 88–93.
- [33] H. Ehrsson, A. Fagergren, R. Johansson, H. Forssberg, Evidence for the involvement of the posterior parietal cortex in coordination of fingertip forces for grasp stability in manipulation, *Journal of Neurophysiology* 90 (2003) 2978–2986.
- [34] S. Imazu, T. Sugio, S. Tanaka, T. Inui, Differences between actual and imagined usage of chopsticks: an fmri study, *Cortex* 43 (3) (2007) 301–307.
- [35] K. Ogawa, T. Inui, Lateralization of the posterior parietal cortex for internal monitoring of self- versus externally generated movements, *Journal of Cognitive Neuroscience* 19 (2007) 1827–1835.
- [36] A. Billard, M. J. Mataric, Learning human arm movements by imitation:: Evaluation of a biologically inspired connectionist architecture, *Robotics and Autonomous Systems* 37 (2-3) (2001) 145–160.
- [37] Y. Kuniyoshi, M. Inaba, H. Inoue, Learning by watching: Extracting reusable task knowledge from visual observation of human performance, *IEEE. Trans. on Robotics and Automation* 10 (6) (1994) 799–822.
- [38] S. Calinon, F. Guenter, A. Billard, On learning, representing and generalizing a task in a humanoid robot, *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special issue on robot learning by observation, demonstration and imitation* 37 (2).
- [39] K. Dautenhahn, C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, Cambridge, MA, 2002.
- [40] M. A., I. H., Representation of bodily self in the multimodal parieto-premotor network, in: S. Funahashi (Ed.), *Representation and Brain*, Springer, 2007, pp. 151–176.
- [41] J. Saarinen, T. Kohonen, Self-organized formation of colour maps in a model cortex, *Perception* 14 (6) (1985) 711–719.



- [42] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning internal representations by error propagation, Vol. 1 of *Parallel distributed processing: explorations in the microstructure of cognition*, MIT Press, Cambridge, 1986.
- [43] R. Shepard, L. Cooper, *Mental images and their transformations*, MIT Press, Cambridge, Mass., 1982.
- [44] R. P. N. Rao, A. P. Shon, A. N. Meltzoff, A bayesian model of imitation in infants and robots, in: K. Dautenhahn, C. Nehaniv (Eds.), *Imitation and Social Learning in Robots, Humans, and Animals: Behavioural, Social and Communicative Dimensions*, Cambridge University Press, Cambridge, 2004.
- [45] T. Inamura, Y. Nakamura, I. Toshima, Embodied symbol emergence based on mimesis theory, *International Journal of Robotics Research* 23 (4) (2004) 363–377.
- [46] J. Namikawa, J. Tani, A model for learning to segment temporal sequences, utilizing a mixture of rnn experts together with adaptive variance, *Neural Networks* 21 (2008) 1466–1475.