Psychological Research (2009) 73:545–558
DOI 10.1007/s00426-009-0236-0

ORIGINAL ARTICLE

Development of hierarchical structures for actions and motor imagery: a constructivist view from synthetic neuro-robotics study

Ryunosuke Nishimoto · Jun Tani

Received: 8 October 2008 / Accepted: 8 January 2009 / Published online: 8 April 2009

Abstract The current paper shows a neuro-robotics experiment on developmental learning of goal-directed actions. The robot was trained to predict visuo-proprioceptive flow of achieving a set of goal-directed behaviors through iterative tutor training processes. The learning was conducted by employing a dynamic neural network model which is characterized by their multiple time-scale dynamics. The experimental results showed that functional hierarchical structures emerge through stages of developments where behavior primitives are generated in earlier stages and their sequences of achieving goals appear in later stages. It was also observed that motor imagery is generated in earlier stages compared to actual behaviors. Our claim that manipulatable inner representation should emerge through the sensory–motor interactions is corresponded to Piaget's constructivist view.

Introduction

How can humans as well as artificial agents acquire diverse skills for goal-directed actions in a flexible, fluent, robust, and context-dependent manner? As a common sense, we know that human infants develop such skills by having rich sensory–motor interaction experiences day by day. Then, question is what are the underlying developmental principles of transforming such experiences to skills?

R. Nishimoto · J. Tani (✉)
RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
e-mail: tani@brain.riken.go.jp

R. Nishimoto
e-mail: ryu@brain.riken.go.jp

Our group has investigated possible neuronal mechanisms of learning goal-directed skilled actions by conducting synthetic neuro-robotics experiments and by analyzing their results with utilizing the dynamical systems framework (Beer, 1995; Schoner & Kelso, 1988; Smith & Thelen, 1994; Tani & Fukumura, 1994). Especially, the studies have focused on the possibility that the anticipation learning paradigm (Butz, Sigaud, Pezzulo, & Baldassarre, 2007; Jordan & Rumelhart, 1992; Pezzulo, 2008; Tani, 1996; Wolpert & Kawato, 1998) embedded in neuro-dynamics with rich sensory–motor interactions could result in acquiring generalized dynamic structures for performing a set of desired goal-directed actions (Tani, Ito, & Sugita, 2004; Tani, Nishimoto, & Paine, 2008b). The essential idea is that anticipatory learning of direct sensory feedbacks associated with each intended action would result in self-organization of "internal reality" (Butz, 2008) those are truly grounded to the actual experiences of the agents. And this idea is quite analogous to Piaget's theories on developmental psychology (Piaget, 1954) which consider that any representations which children might have should have developed through sensory–motor level environmental interactions accompanied by goal-directed actions.

In general views, human skilled actions look too diverse and too complex to be constructed by single level mechanisms. They might require certain hierarchy. The motor schemata theory by Arbib (1981) postulates that a complex goal-directed action can be decomposed into sequence of reusable behavior primitives. On other way around, the theory says that diverse actions can be generated by means of the higher level combining the reusable primitives stored in the lower level in a compositional way. If this type of hierarchical mechanism actually accounts for human skilled actions, essential questions might be how the levels can be organized and then how each level can be developed with interacting with other levels.

The anticipation mechanism enables various mental processes of covert behaviors such as motor imagery or motor planning (Decety, 1996; Jeannerod, 1994; Tani, 1996; Ziemke, Jirenhed, & Hesslow, 2005). Motor imagery is considered as a process during which an individual mentally simulates a given action. This process involves with "the first person perspective" in which an individual feels herself/himself performing the action (Decety, 1996). Tani (1996) showed neuro-dynamics modeling of mental simulation in which cognitive agents become able to engage in simulated interaction with the environment through explorative learning of the environment with utilizing forward models (Jordan & Rumelhart, 1992; Kawato, Maeda, Uno, & Suzuki, 1990) which serve to prepare motor programs before initiating physical actions. Hesslow (2002) also showed a similar idea in terms of mental simulation hypothesis recently. Jeannerod (1994) suggests that lookahead prediction capability by means of motor imagery enables immediate modification of motor program in cases of sudden unexpected changes in environment. It has been also reported that training by mental simulation of actions can develop their performances as comparative to that by physical actions (Feltz & Landers, 1983; Vogt, 1995). If motor imagery is an indispensable competency in accommodating adequate actions, one interesting question might be how motor imagery can be developed as compared to physical actions.

Recently, we proposed a novel neural network model so-called the sensory forward model which utilizes distributed representation scheme embedding multiple goal-directed behaviors in a single neural network model (Nishimoto, Namikawa, & Tani, 2008). The sensory forward model (Nishimoto et al., 2008) anticipates coming sensation of visuo-proprioceptive (VP) state (the egocentric visual state and the body posture state) based on specified goal by means of forward dynamics of continuous-time recurrent neural network (CTRNN) model (Doya & Yoshizawa, 1989). By utilizing the initial sensitivity characteristics of the nonlinear neuro-dynamics, different anticipatory trajectories of VP patterns are learned to be generated depending on the initial states given as the desired goals.

In this model, it is assumed that the anticipation of visuo-proprioceptive flow is performed in inferior parietal lobe (IPL) by which the visual state and the proprioceptive state of predicted are fed-back to visual cortex and somatosensory cortex, respectively. Furthermore it is assumed that the predicted body posture state in terms of the proprioception might be sent to motor cortex as the next step target where necessary motor torques to achieve this target posture is obtained. Although IPL has been considered as a passive integrator of multi-modal perceptions, there are growing evidences (Ehrsson, Fagergren, Johansson, & Forssberg, 2003; Eskandar & Assad, 1999) those support the idea that IPL might involve in positive anticipation of future visuo-proprioceptive state. In addition, recently Imazu, Sugio, Tanaka, and Inui (2007) showed novel evidences from fMRI imaging studies that sensory prediction is conducted in IPL with utilizing the internal model acquired in cerebellum. Therefore, there is a possibility that both IPL and cerebellum contribute to achieve the sensory forward model in real human brains. The target goal information might be given to some IPL neurons as of their initial states from ventral premotor (PMv) of which function has been highlighted by the ideas of mirror neurons (Rizzolatti, Fadiga, Galless, & Fogassi, 1996) or from lateral prefrontal neurons of which roles in generating goal-directed planning are well known (Fuster, 1989).

To scale the learning capability of the original form of the sensory forward model (Nishimoto et al., 2008) we recently proposed a dynamic neural network model consisting of neuron groups with multiple time-scales activation dynamics (Yamashita & Tani, 2008). It was shown that meaningful functional hierarchy can emerge with taking advantages of time-scale differences among the groups (Yamashita & Tani, 2008). The characteristics of self-organization of implicit hierarchy with distributed representation shown in this model contrast with the conventional localist view (Jordan & Jacobs, 1994; Tani & Nolfi, 1999; Wolpert & Kawato, 1998) that assumes explicit local modules and their overt manipulations.

The current paper describes our novel robotics experiments using this architecture with focusing on the aspect of developmental learning of goal-directed skilled actions with human interactive tutoring. The experimental results will clarify the structural relationship among developments of the sensory–motor primitives and their manipulations as well as developments of physical behaviors and motor imagery. Our analysis and discussions will show a possible psychological mechanisms of how manipulatable representations with compositionality could naturally develop solely through sensory–motor experiences in anticipatory behaviors of goal-directed agents.

Model

This section describes how the ideas of the sensory forward model can be implemented in so-called the Multiple Time-scales RNN (MTRNN) (Yamashita & Tani, 2008). Because of presumptions of general readers as well as limited space in the current special issue, the model is described intuitively with abstraction in details. The precise mathematical descriptions should refer to Yamashita and Tani, 2008.

General

The model assumes that a humanoid robot with a simple vision system learns mutiple goal-directed tasks of manipulating an object under tutor supervision. The goal for each task trajectory is provided by the experimenter to the robot by setting the initial state of some neurons in the network model employed. Inputs to the system are the arm joints encoder readings $\hat{p}_t$ (eight dimensional vector of normalized) and two dimensional vector of the camera head angle $\hat{v}_t$ representing object position (Fig. 1). The camera head is programmed to target a red point marked on the frontal surface of the object. Those two modalities of inputs are sparsely encoded in the form of a population coding by using the topology preserving map (TPM) where $\hat{P}_t$ proprioceptive state and $\hat{V}_t$ vision state are obtained. This topology preserving sparse encoding of visuo-proprioceptive (VP) trajectories, which resembles information processing in the primary sensory cortices such as V1 and S1, reduced overlap between VP sequences and improved the learning capacity of the MTRNN.

Based on the current $\hat{p}_t$ and $\hat{v}_t$ the system generate predictions of proprioception p_{t+1} and the vision sense v_{t+1} for the next time step. This prediction of the proprioception p_{t+1} is sent to the robot in the form of target joint angles and actual joint movements are made by a built-in PID controller. Changes in the environment, including changes in object position and changes in the actual position of joints, were sent back to the system as sensory feedback.

The main component of the system modeled by the MTRNN receives the current input of VP state and it outputs the prediction of its next step state. The goal for each task trajectory is given as the initial state in terms of the potential states of slow context units at the initial step. The generation of each task trajectory is made possible by the capacity of the RNN to preserve the intentionality

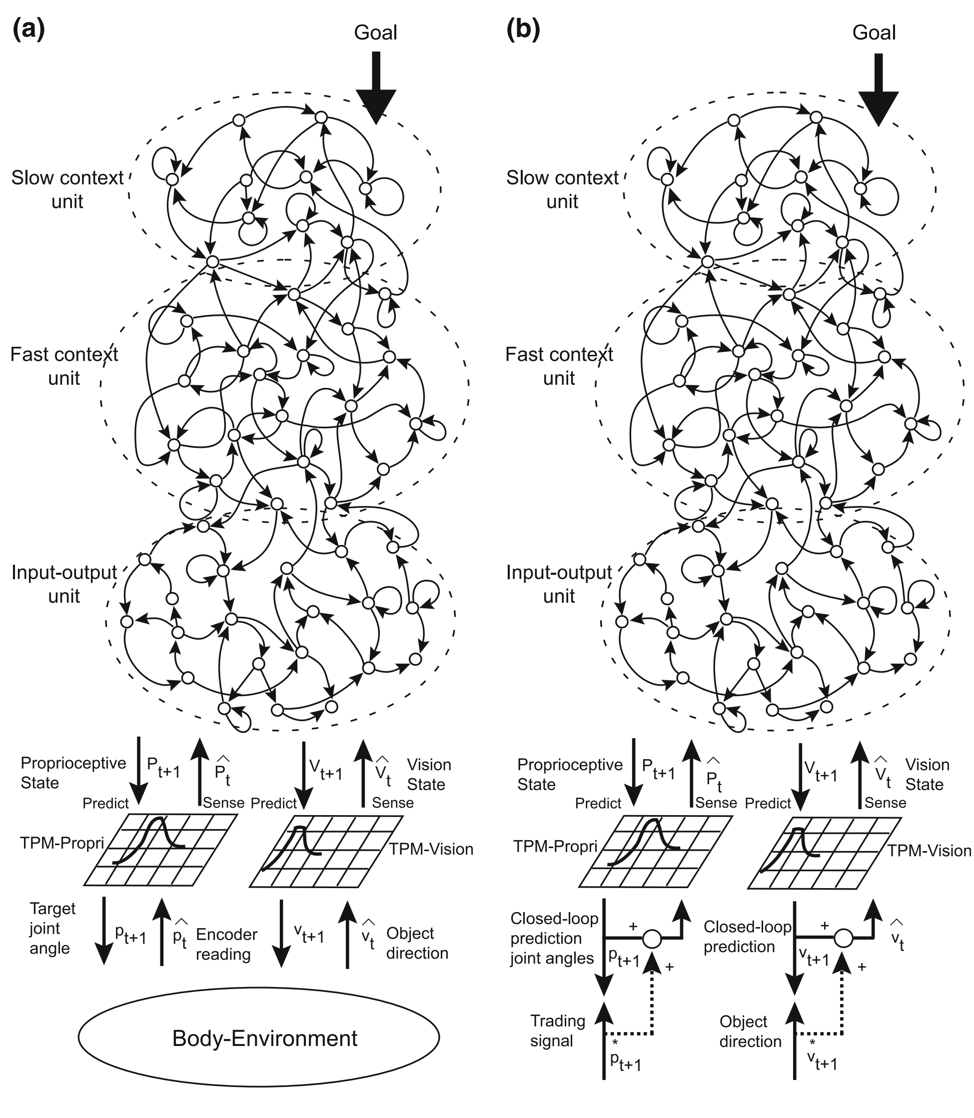

Fig. 1 The MTRNN architecture in the behavior generation mode (**a**) and in the motor imagery and the training mode (**b**)

toward the corresponding goal as the internal dynamics utilizing the slow context units activities.

A conventional firing rate model, in which each unit's activity represents the average firing rate over a group of neurons, is used to model neurons in the MTRNN. In addition, every unit's membrane potential is assumed to be influenced not only by current synaptic inputs, but also by their previous state. In the MTRNN, each neural unit activation is defined with continuous-time dynamics (Doya & Yoshizawa, 1989) of which characteristic is described by the following differential equation, which uses a parameter τ referred to as the time constant:

$$\tau_i \frac{\mathrm{du_{i,t}}}{\mathrm{d}t} = -u_{i,t} + \sum_j w_{ij} a_{j,t} \tag{1}$$

where $u_{i,t}$ is the membrane potential of each ith neuronal unit at time step t, $a_{j,t}$ is an activation of jth unit and w_{ij} is synaptic weight from the jth unit to the ith unit. The current activation state of each unit is obtained as a sigmoidal output of its potential. The time constant τ mostly determines the time scale of the unit activation dynamics. When it is set with large values, the dynamics becomes slow and otherwise quick. Some modeling studies (Nishimoto et al., 2008; Nolfi, 2002) have shown that τ affects strength of context-dependent memory effect in adaptive behavior.

The network that was used in the current model consisted of input–output and non-input–output units, the latter referred to as context units. Context units were divided into two groups based on the value of time constant τ. The first group consisted of fast context units with small time constant ($\tau = 5$) whose activity changed quickly, whereas the second group consisted of slow context unit with a large time constant ($\tau = 70$) whose activity, in contrast, changed much more slowly. Among the input–output units, units corresponding to proprioception and units corresponding to vision are not connected to each other directly. The slow context units and the fast context units are fully connected each other and the input–output units and the fast context units do so as well while the slow context units and the input–output units are not directly connected.

Training

In order to obtain a teaching signal, the experimenter guides both hands of the robot along the trajectory of the goal action. As the robot hands are guided along the trajectory, the sensed VP sequences are recorded, and they were used as teaching sequences. For each behavior task, the object was located in three different positions (center position, 2 cm left of the center and 2 cm right of the center). The objective of learning was to find optimal values of connective weights minimizing the error between teaching sequences and model outputs. At the beginning of training, synaptic weights of the network were set randomly, resulting in the network generating random sequences. Synaptic weights were modified based on the error between teaching signals and generated sequences. After many repetitions of this process, the error between teaching sequences and model outputs eventually reached a minimum level.

This training process is conducted in an off-line manner in the sense that all teaching sequences gathered at each tutoring session are assumed to be stored in a short-term memory (this part is out of the scope) and they are utilized as teacher sequences for consolidation learning of the sensory-forward model assumed in IPL. The tutoring session with gathering new training sequences will be iterated in the course of development. At each training process, lookahead prediction of the VP sequence is generated by means of so-called closed-loop operations (Fig. 1b) in which the current prediction of the VP state are used as input for the next time step. Then, the error between the teacher sequences and the lookahead sequences of imagery are taken by which error-driven training of the network is conducted. The purpose for employing this closed-loop operation in training is to enhance generation of stable dynamic structures of the network by minimizing the error integrated during long steps of lookahead prediction. Our preliminary trials indicated that conventional training scheme of utilizing one-step prediction instead of lookahead one has difficulty in acquiring stable long-time correlations because the error generated at each step becomes too small.

Utilizing the characteristic of initial sensitivity, the network dynamics is trained to generate multiple behavior sequences through adaptation of the initial states of slow context units. In the proposed model, a network is trained by means of supervised learning using teaching sequences obtained through tutoring by the experimenter. The conventional back-propagation through time (BPTT) algorithm (Rumelhart, Hinton, & Williams, 1986) is used for adaptation of both connective weights common to all sequences and the initial state of slow context units for each sequence (Nishimoto et al., 2008). (The initial states of fast context units are not adapted but set as neutral).

Action generation in physical environment and motor imagery

Through the training process, the network learns to predict the VP inputs for the next time step. The prediction of proprioceptive state provides the target joint angles to the robot controller which enables the robot to generate movements.

Moreover, by using the prediction of VP feedback as input to the next time step (closed loop operation), the network can be able to autonomously generate VP trajectories without producing actual movements. This process of closed loop generation may correspond to motor imagery in terms of mental simulation of actions (Decety, 1996; Jeannerod, 1994; Tani, 1996). It is noted that the motor imagery in the current paper is defined as image of inseparable coming flows of kinesthetic one and egocentric visual one in terms of VP trajectory.

Setup of humanoid robot experiments

A small humanoid robot was used in the role of a physical body interacting with actual environment. A workbench was set up in front of the robot, and a cubic object (approximately 9 × 9 × 9 cm) placed on the workbench served as the target object for manipulations. The robot task is to learn to generate three different task behaviors. The goal of each task behavior is to generate a different sequence of behavior primitives of manipulating the object. All task behaviors start from the home position and end with going back to the same position(Fig. 2).

In the task-1, with starting from the home position, the both hands grasp the object, move the object up and down (UD) for four times, do it for left and right (LR) for four times and go back to the home position (BH). In the task-2, the object is moved forward and backward (FB) for four times and then it is touched by left and right hands bilaterally (TchLR) for four times and finally BH. In the task-3, the robot repeats grasping and releasing the object by both hands (BG) for four times and then BH. A tutor teaches the robot with these three task behaviors in three tutoring sessions with changing the object position three times from the center position to the left and to the right for each task behavior. In the first session, the robot guidance is conducted by disabling active movements of the robot by setting the motor control gain to zero because the networks are not yet effective with the randomly set initial synaptic weight values. In the second and third sessions, the tutoring is conducted interactively by enabling active movements of the robot with the control gain set to 20% of its normal operation value. The network is trained off-line by using all tutoring sequence data obtained at each session. The network consists of 144 proprioceptive units, 36 vision units, 30 fast context units and 20 slow context units.

During these three training sessions some learning related parameters are tuned in order to realize smooth training processes. One parameter is the so-called closed-loop ratio *CLr*. The current step VP inputs are weighted sum of the predicted one in the previous step and the target one in the current step by *CLr*, whereas *CLr* set as 1.0 means a complete closed-loop operation (lookahead prediction) and that of 0.0 does a complete open-loop one (one-step prediction). Although the complete closed-loop operation enhances the learning process substantially because of larger integrated error as have been described previously, this could also break down the learning process by accompanying sudden catastrophic changes in synaptic weights. On other hand, the complete open-loop operation cannot result in rigid structural learning because of quite small amount of the error. Because it is better to relax the training in the early period, *CLr* is set with a smaller value in the earlier sessions than later ones in terms of annealing. In fact, it was observed in our preliminary experiments that the training did not converge well with *CLr* set with 0.9 in the first session regardless of amount of the training epochs. σ^I and σ^F are other parameters to be changed for the self-organization of the TPM. The idea is that the receptive field in the TPM should be changed from a wide one with large σ of σ^I to a sharp one of small σ^F for the

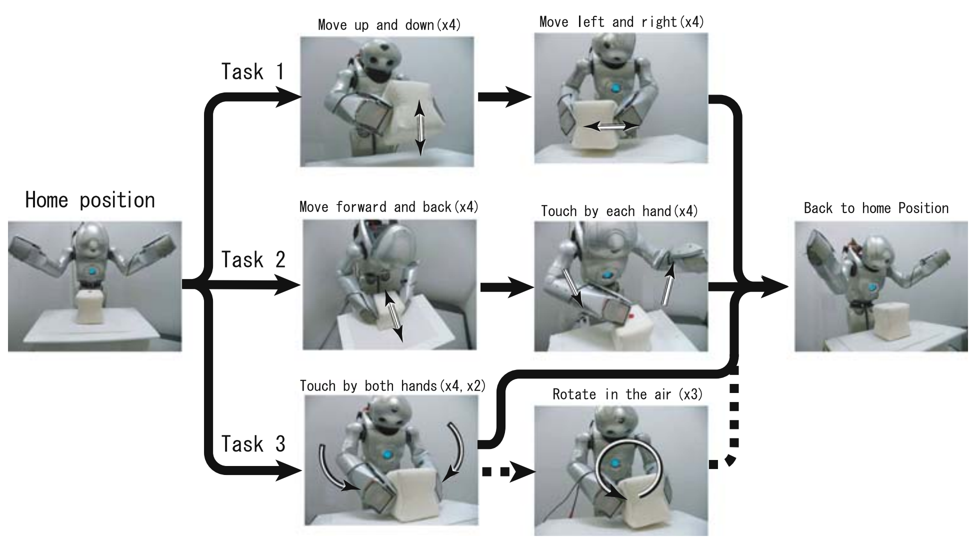

Fig. 2 Three task behaviors tutored to the robot. After the 3rd session, the task-3 is modified as illustrated by *dot lines*

relaxation reason. More descriptions about σ for the TPM self-organization should be referred to the appendix. The setting of all these parameters for each tutoring session is shown in Table 1.

After the basic tutoring of three sessions, the task-3 is modified with introducing a novel behavior primitive which is to rotate the object in the air (RO) by both hands. In the session 4 and 5 of the task 3, BG is repeated two times followed by three times repetitions of RO. This additional tutoring is conducted to examine the capability of the network to incrementally learn novel patterns. In the session 4, the training parameters are once relaxed in order to minimize the interference between the previously learned contents and the new one (see the 4th and 5th session in Table 1. It is noted that the interference could occur not only in cases of introducing novel primitives but also for novel sequential combinations of them, because this requires fine adjustments in both of the lower and higher levels to achieve end-to-end smooth connections between the primitives.

Table 1 The parameter setting for each training session

	Session 1	Session 2	Session 3	Session 4	Session 5
σ_{PropI}	300	1.5	0.75	0.75	0.75
σ_{PropF}	1.5	0.75	0.375	0.375	0.375
$\sigma_{VisionI}$	150	0.75	0.375	0.375	0.375
$\sigma_{VisionF}$	0.75	0.375	0.1875	0.1875	0.1875
CLr	0.6	0.8	0.9	0.6	0.9

Results

Overall task performances in the end of development

The developmental tutoring experiment was repeated twice with setting the initial synaptic weights of the networks as randomized. Figure 3 shows how the robot behaviors were generated in the test run after the five sessions of tutoring in one developmental case. Plots are shown for the VP trajectories (sequences of two representative arm joint angles denoted as "Prop #" and two camera head angles of normalized denoted as "Vision #") in the tutoring in the top row, the actually generated one in the second row and the fast and slow context activations represented by the first four principal components denoted as "PC #" after their principal component analysis (PCA) in the third and the forth row, respectively, for all three tasks. It is observed that the actual VP trajectories are exactly reconstructed from the tutoring ones for all the tasks. Actually, the robot was successful in performing all tasks with the object

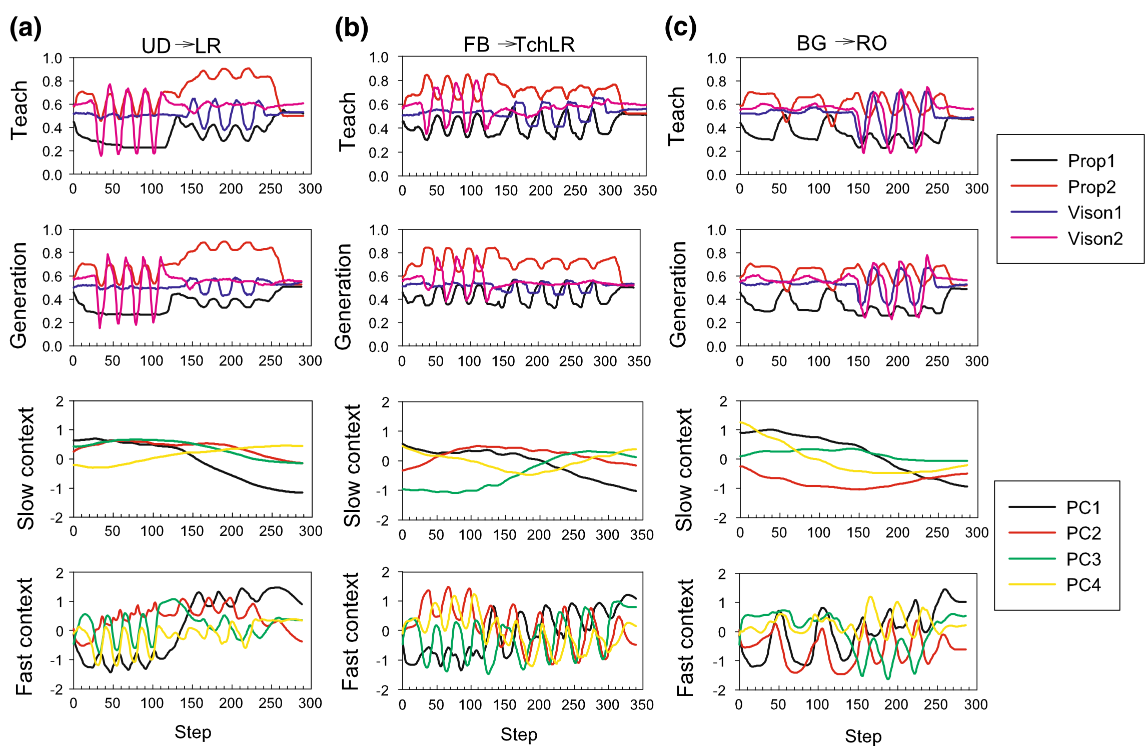

Fig. 3 VP trajectories (two normalized joint angles denoted by Prop1 and Prop2 and camara head direction denoted by Vision1 and Vision2) in tutoring and in actual generation accompanied with fast and slow context profiles with PCA denoted by PC1, PC2 and PC3. **a** Task 1, **b** Task 2, **c** Task 3

position varied within the range of tutored after the five tutoring sessions. The profiles of the fast context activations and those of the slow ones can be contrasted. The fast ones mostly synchronize with the VP trajectories while the slow one shows smoothly changing trajectory starting from different initial state of self-determined for each task. It is observed that the slow context profiles abruptly change when the cyclic pattern in the VP trajectories shift from one primitive to another. These observation suggest that each exact pattern of the primitives is embedded in the fast context activation dynamics while each macro scenario of sequencing of the primitives is embedded in the slow context one. This result accords with the one in (Yamashita & Tani, 2008).

Development processes

Now, the development process is closely examined as the main focus of the current paper. Figure 4 shows one developmental case of the task-1 with the object located in the center from session 1 to session 3 before the novel task behavior is introduced in the task-3. Plots are shown for the VP trajectories of tutoring in the left, motor imagery in the middle, and actual robot generation in the right. The slow context units profiles in the motor imagery and the actual behavior are plotted for their first four principal components after the PCA. It is noted that the tutoring trajectories in the session 1 is quite distorted. The tutoring patterns of UD in the first half and LR in the second half are not regular cycles. This is a typical case when cyclic patterns are tutored to robots without using metronome-like devices. However, it can be seen that the cyclic patterns in the tutoring become much more regular as the session proceeds.

One interesting observation is that the motor imagery patterns develop faster than the actual ones over these three sessions. In the session 1, the cyclic pattern of UD is successfully generated (but not for LR) in the motor imagery while neither UD nor LR are yet generated in the actual behavior generation. It is noted that the cyclic pattern of UD is more regular than the tutored one in the session 1. In the actual behavior generation, the robot hands touched the object but not accurate enough to grasp and hold it up and after the failure the movements were frozen. One interesting observation was obtained by conducting an extra experiment using fake visual feedback. In this extra experiment, the tutor grasped the object and moved it up and down immediately after the robot touched the object. It turned out that the robot hands moved up and down correctly following the object movement of perceived. It can be understood that this arm movement was generated by means of the entrainment with the fake visual feedback. The same phenomena had been observed in the case using a modular network model of the mixture of RNN experts (Tani, Nishimoto, Namikawa, & Ito, 2008a).

In the session 2, both UD and LR cyclic patterns are generated in the correct sequence in motor imagery while only UD pattern is generated which cannot be shifted to LR pattern in the actual behavior. This can be explained by an observation that the slow context profile in the motor imagery dynamically changes around 160 steps while that of actual behavior does not show any significant changes around this transition period. It is considered that the dynamics of the slow context units is not strong enough to generate the shifting in the actual behavior interacting with noisy environment. This consideration is supported by the fact that the robot could actually make the shift when the tutor assisted the robot to do so by guiding the arm trajectories with force in the transition period.

In the session 3, both UD and LR patterns are successfully generated both in the motor imagery and in the actual behavior generation. It was, however, observed in limited cases that counting of repetition times of primitives (as like UD four times) could go wrong within the range of plus or minus one time probably by perturbed by noise during physical execution of actions. An interesting observation here is that even when the counting goes wrong, smooth transition from one primitive to another is still preserved e.g., moving object left and right always follows immediately after the object is once placed on the table. The transition never takes place by cutting through in the middle of on-going primitives. This observation implies firstly that counting in the higher level is more like implicit and analogical process rather than explicit and logical one and secondly that the lower level is successful in organizing fluidity in connecting primitives which could be expressed by Luria's (1973) metaphor of "kinetic melody".

Analyses

In this section, more detailed analyses are shown for examining the observed developmental processes. Figure 5 shows how the success rate in metal simulation and in actual behavior generation change in average of all three task behaviors for the two developmental cases. Here, the success rate is defined as rate of how many primitive events can be successfully generated as in the trained order in mental simulation and actual behavior. For example, the number of primitive events in the task-1 is counted as 9 with 4 UDs, 4 LRs and 1BH. The number is counted by looking at the robot behavior for actual behavior and by examining plots of VP trajectories generated compared with the trained one for mental simulation. In Fig. 5, it can be seen that the success rate of mental simulation is higher than the one of actual behavior at least the first three

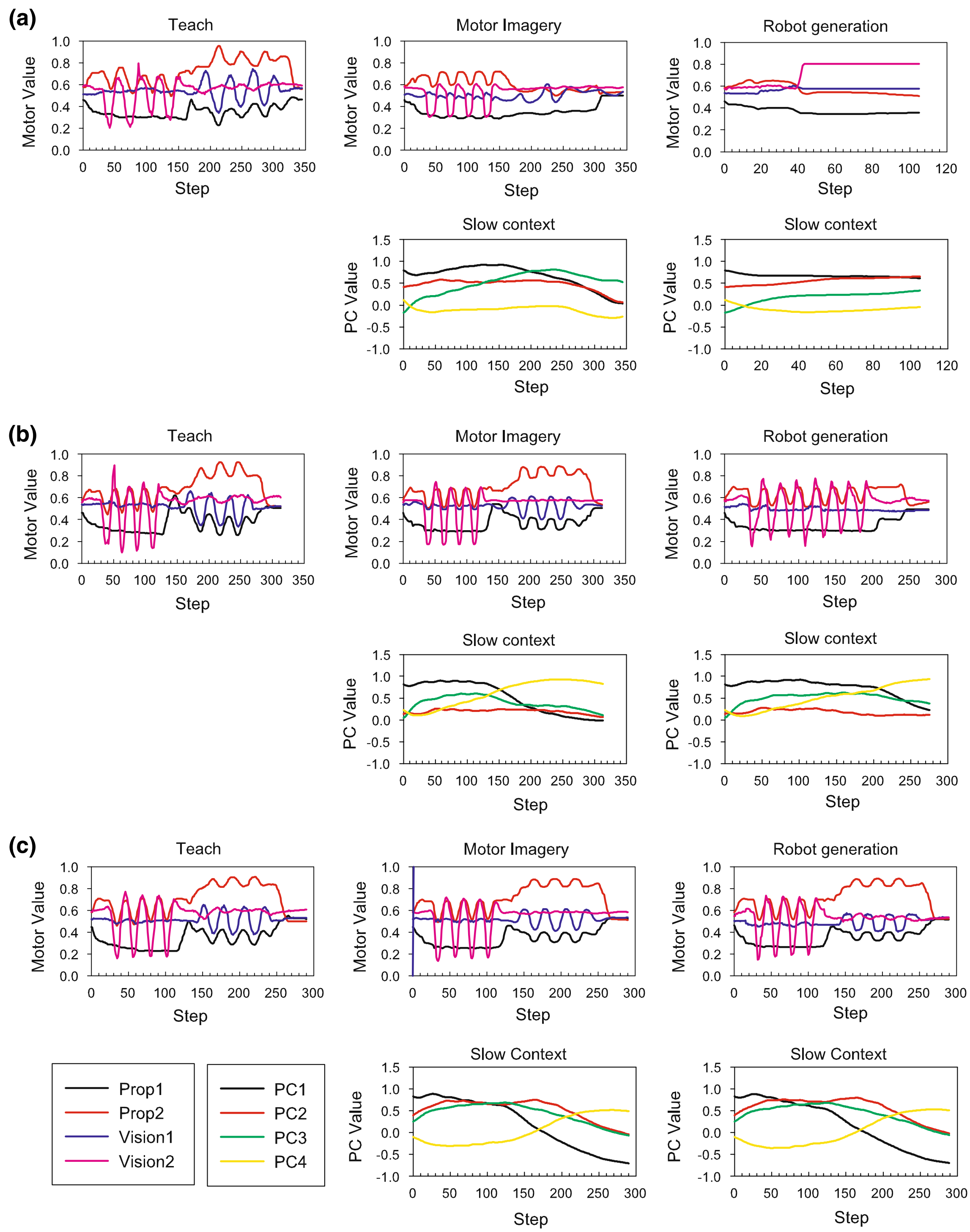

Fig. 4 Development of task-1 for the initial 3 sessions with VP trajectories of tutoring, motor imagery and actual generation accompanied with slow context profiles by PCA. **a** Session 1, **b** Session 2, **c** Session 3

sessions in both developmental cases. It is observed that the success rate becomes 1.0 as perfect after 3 sessions of tutoring for both mental simulation and actual behavior. Then the rate slightly decreases in the session 4 when the novel task behavior is introduced in the task-3. It, however, goes back to nearly 1.0 in the session 5.

Figure 6 shows the success rate for each task behavior in actual behavior in developmental cases. It can be observed

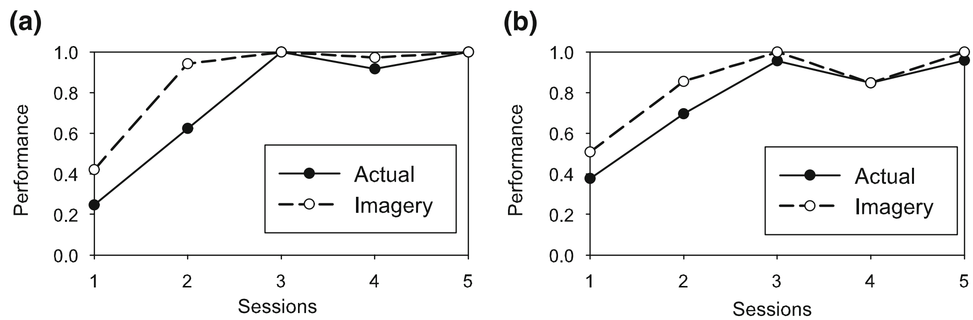

Fig. 5 The developments of success rate averaged over three task behaviors in motor imagery denoted by "Imagery" and in actual behavior denoted by "Actual" for 2 case runs. **a** Case 1, **b** Case 2

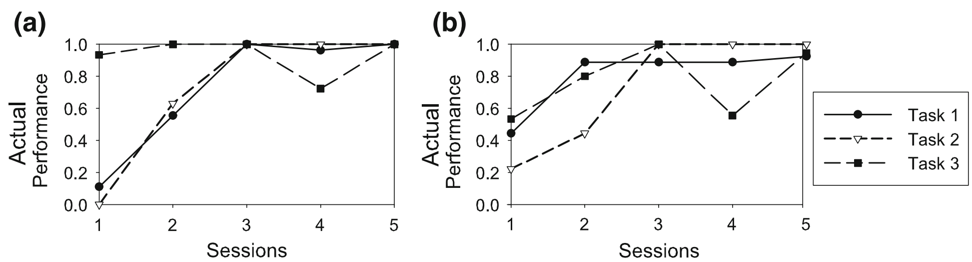

Fig. 6 The developments of success rate of each task behavior denoted as "task1, task2 and task3" in actual behavior for 2 case runs. **a** Case 1, **b** Case 2

that the success rates of task-1 and task-2 stay near 1.0 after session 3 to the end while that of task-3 once decreases in session 4 when the novel behavior primitive RO is introduced and it reaches to 1.0 in the end. This result indicates that introduction of a novel behavior primitive in a task behavior does not affect the performances in other task behaviors unless they share the same behavior primitives. It was also observed that the behavior primitive of BG, which was followed by RO, was not distorted in the session 4 in both development cases. Only RO was immature. This means that once acquired primitives can be utilized in generating different sequential combinations of the primitives. Such recombination capability for primitives was also shown in our prior study (Yamashita & Tani, 2008).

Figure 7 illustrates how the encoding of the basic primitives by the fast context units develop during the first three sessions. The trajectory of the fast context units during each basic primitive pattern in mental simulation is plotted as a phase diagram using the first and the second principal components. Three colors in the plots denote three different object positions cases. It is observed that there are no regularities in the patterns shown for the session 1 except the BG case in which each trajectory shows a near cyclic pattern that is shifted with the position difference. In the session 2, such near cyclic patterns appear for all basic primitives. Finally, we see that the shapes of the patterns in the session 3 are mostly similar to the ones in the session 2. These results imply that basic primitives begin to be embedded in pseudo attractor of limit cycling by the fast context units with achieving the object position generalization in the second session. This development by fast context units seems to converge mostly in the second session.

Discussion

Summary of the robot experiments

Now, the robotics experiment results are summarized with qualitative discussions. It was shown that the developmental learning processes of multiple goal-directed actions were successfully converged after several sessions of the teacher tutoring. The developmental process can be categorized in some stages. In the first stage, which mostly corresponds to the session 1, no tasks are completed where most of behavior primitives in actual generation are premature. In the second stage corresponding to the session 2, most of behavior primitives can be actually generated although their sequencing is not yet completed. In the third stage corresponding to on and after the session 3, all tasks are successfully generated with correct sequencing of the primitives. From this observation, it can be said that there is an order in the formations of different levels of functionality. The level for behavior primitives is generated by the 2nd stage while the level for sequencing the primitives does by the 3rd stage. It is natural that the primitive level as the lower level is organized earlier and then the level for the sequencing as the higher level does later based on the prior formation of the lower level.

However, one interesting remark is that there is a time lag between the period of becoming able to generate motor imagery and actual behavior. The motor imagery is generated earlier than the actual behavior as it was seen that the motor imagery for all tasks are nearly completed by the session 2 as compared to the session 3 by the actual ones. This issue will be revisited with some psychological

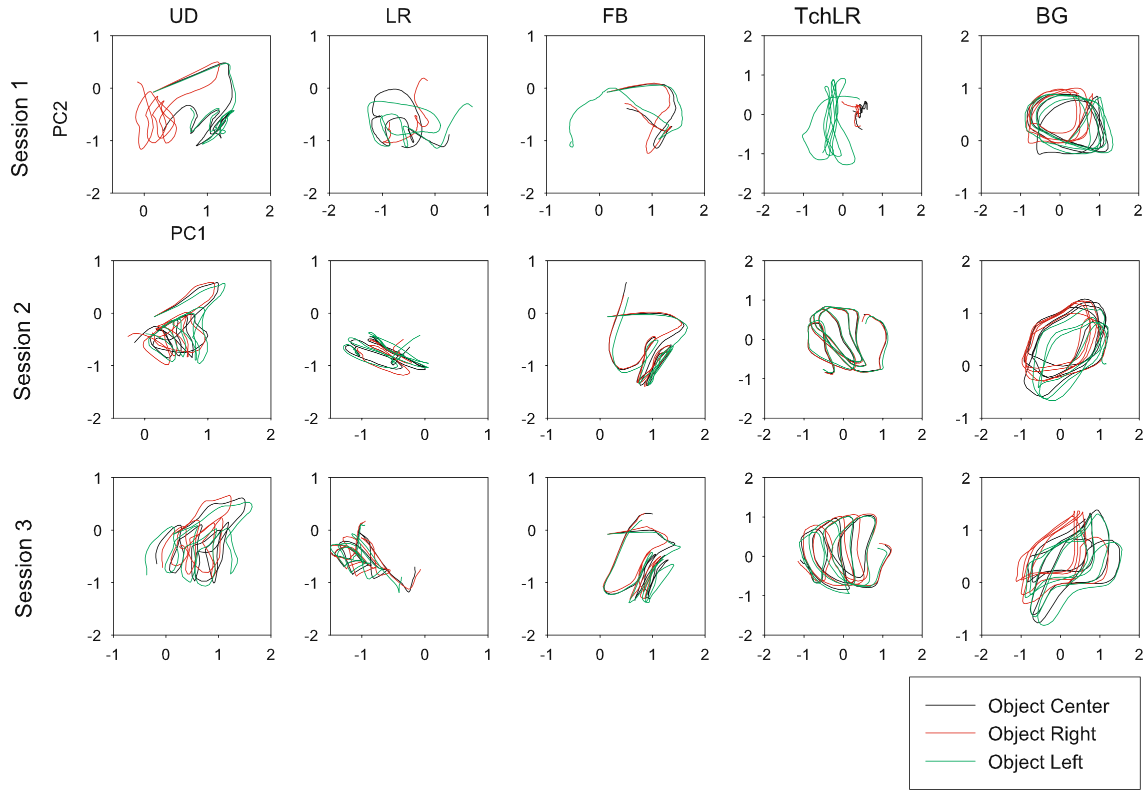

Fig. 7 The development of encoding of basic primitives by the fast context units with PCA

considerations later in this section. Another remark is that when a new task which composed of a prior-trained primitive and a novel one was introduced midway, the introduction affects the overall task performances only slightly. Although regeneration of the novel primitive is premature initially, the prior-trained primitive is well adopted in this new task and also performances of other tasks are intact.

Correspondences to psychology of development and learning

The above mentioned qualitative observation in our robotics experiment could correspond to some psychological observations and theories for development and learning with abstraction. Among them Piaget's constructivist accounts for infant developments might be most relevant. Piaget considered that if infants can have representations, they should self-organize through the dynamic interactions between subject and object. The dynamic interactions should involve with the one in sensory–motor level accompanied with goal-directed intentionality about the object. Then, operative and figurative aspects of intelligence should emerge as the results of self-organization through such dynamic interactions. There are two core concepts those compose the Piaget's theory. One is assimilation and the other is accommodation. Assimilation is a process that existing scheme of subject is exploited to establish structural coupling with object. On other hand, accommodation is an adaptive process to modulate the scheme to establish another structural coupling with object.

If we look at our experiments, it is understood that scheme in Piaget's theory may correspond to a set of behavior primitives embedded in the fast context network. Depending on the top–down signal conveying the task goal information flowing from the upstream slow context network, different dynamic structures of the primitives are adopted which may explain dynamic mechanism of assimilation. These behavior primitives are actually the products of the neuronal self-organization with having rich sensory–motor interactions through iterative tutoring. This may account for accommodation. The case of introducing a new task behavior in the session 4 could be interpreted that both assimilation and accommodation occur because the pre-acquired primitive is utilized in the novel task while a novel behavior primitive is additionally self-organized. The fact that six different behavior primitives were compositionally manipulated to generate both actual behaviors and motor imagery for achieving multiple goals in the end of the developmental tutoring could be interpreted that certain operational representations are finally appeared through the long-term self-organization process. It is, however, argued that the operational representations appeared in this stage is not just compositional, as if composed of a set of discrete symbols, but "organically" composed (Tani et al., 2008a, 2008b) by capturing fluid and contextual nature of human skilled behaviors in neuronal circuits of analog dynamical systems. This argument is supported by the current observations of various phenomena including implicit and analogical counting in repeating primitives and smooth transitions in the primitive sequences.

It is noted that local representation scheme as like hierarchically gated modular networks (Haruno, Wolpert, & Kawato, 2003; Tani & Nolfi, 1999) can also exhibit above mentioned properties of "organic compositionality" (Tani et al., 2008a, 2008b). Actually, our group has conducted a similar developmental learning experiment using the hierarchically gated CTRNN (Tani et al., 2008a, 2008b) where similar phenomena were observed. However, it was found that this localist scheme has more difficulty in tuning parameters related to gating dynamics compared to the current scheme. It was severely difficult to increase number of trained primitives up to the one in the current experiment case because of "near-miss" problems (Tani et al., 2008a, 2008b) in selecting best match modules for the current pattern. Although this near-miss problem could be quite improved by introducing an additional parameter control scheme so-called the adaptive variance in the gating dynamics (Namikawa & Tani, 2008), the whole system becomes much more complex with more parameters. It is speculated that the drawback of the local representation scheme might be originated from its inherent explicitness in operations for segmenting sensory–motor flow into primitives and manipulating them into desired sequences and also in representation of utilizing segregated modules and levels. Such explicitness might hamper natural processes of self-organization in the employed network model.

Our approach is also parallel to the ones by so-called the neo-Piagetian especially who attempt to explain the infant development as time-development of complex systems (Smith & Thelen, 1994). Smith and Thelen (1994) claim that infant development is better understood as the emergent product of many decentralized and local interactions that occur in real time where coherence among local parts is achieved. Our robotics experiments have been carefully designed such that local interactions can be enhanced in different levels. The MTRNN was designed such that neuronal dynamics can interact with row sensory–motor flow in the continuous space and time domain without introducing any apriori articulation mechanisms. Also, there is no algorithmic operations those act on segregated modules of the higher and the lower levels or independent modules of encoding primitives. All there exist are just a single network where different time-scale dynamics coexist and their interactions result in self-organization of functional hierarchy. Furthermore, the tutoring procedure was designed such that the tutor and the robot can directly interact each other with force. It was observed that not only the robot trajectories develop but also the tutoring ones do across sessions to generate smooth and rhythmic patterns. The direct force level interactions enabled this sort of co-developments between the two sides.

The sensory forward model employed in the current study should be distinguished from the conventional forward model (Kawato et al., 1990; Wolpert & Kawato, 1998). The conventional forward model predicts the resultant future sensation for the current motor commands given. One notorious problem is that the forward model cannot predict all the outcomes of possible motor commands because of combinatorial complexity associated with their high dimensional space. This problem is related to the frame problem (McCarthy, 1963) well known in Artificial Intelligence. It tells that an optimal action can not be determined if infinite number of possible action consequences are examined at each step. Why does this happen? This is because the conventional forward model does not deal with goal-directedness. The conventional forward model attempts to predict consequences of arbitrary action sequences which may not be related to any goal achievements. On other hand, the sensory forward model of our proposal attempts to predict coming sensory flow in the course of achieving each specified goal. Because the sensory forward model learns about only finite number of goal-directed paths through the actual tutoring experiences, it never faces with the combinatorial explosive problems. Indeed, Piaget's advocacy of goal-directedness is right in a sense that the burden of goal-directedness with enactments actually avoids unrealistic combinatorial computations in cognition.

However, there is one major drawback in the current formulation of the sensory-forward model. In the current setting, the goal specified by the initial internal state cannot produce multiple possible trajectories of achieving the same goal because of the deterministic dynamics nature of the sensory-forward model. It might be better to consider that the initial internal state corresponds to action program rather than goal state. This is because the sensory-forward model can generate different trajectories of motor imagery reaching to the same distal goal state provided that each of such trajectories has been learned with attaining specific initial internal state. Then, planning to achieve specified goal states can be formulated by means of searching initial internal states such that distal state in motor imagery generated can match with the specified goal state. Our recent study (Arie, Endo, Arakaki, Sugeno, & Tani, 2009) has examined such trials.

Our experiments showed that motor imagery develop faster than actual behaviors. Does it correspond to any reality in human development and learning? Some contemporary developmental psychologists such as Karmiloff-Smith (1992) and Diamond (1991) claim that mental representation develops very earlier in life, or is even innate. It is said that infants of 2 months old already possess intentionality toward objects to deal with but just cannot reach properly to them because of immaturity in motor control skills. It might be plausible that motor imagery of reaching to objects develops easily if infants happen to reach to the objects of their interests by motor bubbling and such episode is reinforced with joys. However, the actual motor acts on objects such as touching or grasping them are far more difficult because they involve with precise arm controls of making physical contacts to the objects, as had been shown in our robotics experiments. Because the generation of motor imagery, on other hand, do not require such fineness, it could be achieved earlier. Flanagan, Vetter, Johansson, and Wolpert (2003) showed evidences that human subjects learn to predict sensory feedback faster than motor control in their psychophysics experiments on object manipulation under artificial force field. This finding might be related to the current results because the predictability of sensory feedback directly links to motor imagery.

Also it is known that generation of motor imagery has a positive role in consolidating memories (Feltz & Landers, 1983; Jeannerod, 1995; Vogt, 1995), as have been mentioned earlier. The robot training scheme shown in the experiment is analogous to this evidence because in our scheme the network is trained to re-generate the teaching sequences in the closed-loop operation without receiving actual inputs as like rehearsing and this explains why motor imagery develop earlier than the actual one.

The motor imagery by means of lookahead prediction can provide means for on-line monitoring of future perspective. If unexpected external changes happen during physical execution of goal-directed actions, the monitoring by the lookahead prediction can detect a future perspective gap as the error between the originally intended goal state and the currently predicted one. The detected error can be utilized to modify the original goal state to currently possible one by modulating the internal state that carries goal information. This on-line monitoring and the error-driven goal state modulation can be implemented by pairing the future lookahead prediction and the past regression as have been described elsewhere (Ito, Noda, Hoshino, & Tani, 2006; Tani, 2003). The motor imagery plays essential roles in accommodating cognitive behaviors in diverse ways including goal-directed motor planning, on-line monitoring of future perspective and resultant gaol modulation, and enhancements of consolidation learning. Future studies should focus to integrate those different functions systematically in synthetic models.

Robotics synthetic approach

How can the synthetic robotics modeling researches contribute to understanding of human development and learning? It is obviously true that the robotics studies cannot reconstruct complete realities of human developmental processes. The studies, however, could show some interesting analogy with the reality. The conventional psychology can be considered as an attempt to elucidate possible underlying mechanisms in a black box solely through observations and analysis of human behavior data. The approach of computational neuroscience goes to an opposite direction in which detailed neuronal mechanisms are investigated by building anatomically relevant neuronal circuitry models based on neuroscience data of neuron connectivities and cell firing properties but without paying much attentions to behavioral data.

In this aspect, the recent connectionist approach applied to developmental psychology (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett, 1997) is considered to come in the midway of these two extremes. Although the connectionist models mimic biological neuronal circuits, their descriptions are quite abstract with neglecting detailed cell firing properties and anatomical connectivities. However, the connectionist pay much attentions to inputs/outputs type functionality of their networks. They attempt to evaluate the functionality by applying the psychological behavioral data as the inputs/outputs of the network models. This approach could provide certain constraints in elucidating possible brain mechanisms such that the mechanisms should be realized by collective activities of massively parallel elements within networks as like real brains do. This approach should be much better than regarding a brain as a black box and evaluating whatever computational mechanisms which best fit with the behavioral data.

Although our synthetic robotics modeling approach inherits the essential characteristics of the connectionist approach, it goes beyond. The robotics experiments provide one more constraint that is cognitive mechanism should be realized in the structural coupling between subject and environment (Beer, 1995). Additionally, the robotics experiments tend to take a holistic approach of integrating necessary ingredients into one trial model. The robotics experiments should deal with multi-modalities of sensations, motor systems, memory, attentions, anticipation, and learning all together at the same time of which attitude is contrasted to the connectionist one of focusing on a single modality of information processing at each model. For example, although Elman's studies of showing how linguistic competency can develop using Elman network (Elman, 1990) is inspiring, the problem of how linguistic semantics can be grounded might be better understood if dynamic interactions between the linguistic modality and the sensory–motor modality are seriously considered as have been shown in our prior robotics study (Sugita & Tani, 2005). If human development and learning should be looked at with a holistic view, the robotics experiments could provide a nice platform to implement such view.

Summary

The current paper showed a neuro-robotics experiment in which developmental learning processes of goal-directed actions of a robot were examined. The robot was implemented with the MTRNN model which is characterized by co-existences of the slow context dynamics and the fast context dynamics in generating anticipatory behaviors. Through the iterative tutoring of the robot for multiple goal-directed actions, certain structural developmental processes emerged. It was observed that behavior primitives are self-organized in the fast context network part earlier and sequencing of them appear later in the slow context part. It was also observed that motor imagery develop faster than the actual ones. The paper discussed that the robotics synthetic experiment results are quite analogous to Piaget's ideas of the constructivism which emphasis the roles of goal-directed sensory–motor interactions in acquiring operational representations in human development.

Acknowledgment The authors thank Sony Corporation for providing them with a humanoid robot as a research platform. They also thank Dr. Giovanni Pezzulo for providing them useful suggestions during the revision of the paper. The study has been partially supported by a Grant-in-Aid for Scientific Research on Priority Areas "Emergence of Adaptive Motor Function through Interaction between Body, Brain and Environment" from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

Appendix: Topology preserving map

The weight of TPM is updated by using the following equation with the neighboring function h.

$$w_i(t+1) = w_i(t) + h_i(t)[x(t) - m_i(t)] \quad (2)$$

$$h_i = \alpha(t)\exp\left(-\frac{||r_c - r_i||^2}{\sigma(t)}\right) \quad (3)$$

Where $x(t)$ and $m_i(t)$ denote the input vector and the reference vector, respectively. In the neighborhood function the learning rate $\alpha(t)$ and the distribution $\sigma(t)$ are annealed with time in the following time schedule.

$$\sigma(t) = \sigma^I\left(\frac{\sigma^F}{\sigma^I}\right)^{\frac{t}{\text{maxstep}}} \quad (4)$$

$$\alpha(t) = \alpha^I\left(\frac{\alpha^F}{\alpha^I}\right)^{\frac{t}{\text{maxstep}}} \quad (5)$$

Where σ^I represents the initial value of σ and σ^F represents the final value of σ. α^I is the initial learning rate. α^F is the final learning rate.

References

Arbib, M. (1981). Perceptual structures and distributed motor control. In *Handbook of Physiology: The Nervous System, II. Motor Control* (pp. 1448–1480). Cambridge: MIT Press.

Arie, H., Endo, T., Arakaki, T., Sugeno, S., & Tani, J. (2009). Creating novel goal-directed actions at criticality: A neuro-robotic experiment. *New Mathematics and Natural Computation, 5*(1), 307–334.

Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence, 72*(1), 173–215.

Butz, M. V. (2008). How and why the brain lays the foundations for a conscious self. *Constructivist Foundations, 4*(1), 1–14.

Butz, M., Sigaud, O., Pezzulo, G., & Baldassarre, G. (2007). *Anticipatory behavior in adaptive learning systems: From brains to individual and social behavior*. Berlin: Springer.

Decety, J. (1996). Do executed and imagined movements share the same central structures? *Cognitive Brain Research, 3*, 87–93.

Diamond, A. (1991). *Neuropsychological insights into the meaning of object concept development*. Hillsdale: Erlbaum.

Doya, K., & Yoshizawa, S. (1989). Memorizing oscillatory patterns in the analog neuron network. In *Proceedings of 1989 international joint conference on neural networks*, washington, D.C. (pp. I:27–32).

Ehrsson, H., Fagergren, A., Johansson, R., & Forssberg, H. (2003). Evidence for the involvement of the posterior parietal cortex in coordination of fingertip forces for grasp stability in manipulation. *Journal of Neurophysiology, 90*, 2978–2986.

Elman, J. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Elman, J. L., Bates, A. E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1997). *Rethinking innateness a connectionist perspective on development*. Cambridge: MIT Press.

Eskandar, E., & Assad, J. (1999). Dissociation of visual, motor and predictive signals in parietal cortex during visual guidance. *Nature Neuroscience, 2*, 88–93.

Feltz, D. L., & Landers, D. M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *Journal of sport psychology, 5*, 25–57.

Flanagan, J., Vetter, P., Johansson, R., & Wolpert, D. (2003). Prediction precedes control in motor learning. *Current Biology, 13*(2), 146–150.

Fuster, J. (1989). *The Prefrontal Cortex*. New York: Raven Press.

Haruno, M., Wolpert, D., & Kawato, M. (2003). Hierarchical mosaic for movement generation. *International Congress Series, 1250*, 575–590.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Science, 6*(6), 242-247.

Imazu, S., Sugio, T., Tanaka, S., & Inui, T. (2007). Differences between actual and imagined usage of chopsticks: an fMRI study. *Cortex, 43*(3), 301-307.
Ito, M., Noda, K., Hoshino, Y., & Tani, J. (2006). Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Networks, 19*, 323–337.
Jeannerod, M. (1994). The representing brain: neural correlates of motor imitation and imaginary. *Behavioral and Brain Science, 17*, 187–245.
Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia, 33*(11), 1419–1432.
Jordan, M., & Jacobs, R. (1994). Hierarchical mixtures of experts and the EM algorithm. *Neural Computation, 6*(2), 181–214.
Jordan, M., & Rumelhart, D. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16*, 307–354.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge: MIT Press.
Kawato, M., Maeda, Y., Uno, Y., & Suzuki, R. (1990). Trajectory formation of arm movement by cascade neural network model based on minimum torque-change criterion. *Biological Cybernetics, 62*(4), 275–288.
Luria, A. (1973). *The working brain*. London: Penguin Books Ltd.
McCarthy, J. (1963). *Situations, actions and causal laws*. (Stanford Artificial Intelligence Project, Memo2).
Namikawa, J., & Tani, J. (2008). A model for learning to segment temporal sequences, utilizing a mixture of rnn experts together with adaptive variance. *Neural Networks, 21*, 1466–1475.
Nishimoto, R., Namikawa, J., & Tani, J. (2008). Learning multiple goal-directed actions through self-organization of a dynamic neural network model: A humanoid robot experiment. *Adaptive Behavior, 16*, 166–181.
Nolfi, S. (2002). Evolving robots able to self-localize in the environment: The importance of viewing cognition as the result of processes occurring at different time scales. *Connection Science, 14*(3), 231–244.
Pezzulo, G. (2008). Coordinating with the future: The anticipatory nature of representation. *Minds and Machines, 18*, 179–225.
Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
Rizzolatti, G., Fadiga, L., Galless, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131–141.
Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing* (pp. 318–362). Cambridge: MIT Press.
Schoner, S., & Kelso, S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science, 239*, 1513–1519.
Smith, L., & Thelen, E. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge: MIT Press.
Sugita, Y., & Tani, J. (2005). Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adaptive Behavior, 13*(3), 33–51.
Tani, J. (1996). Model-Based Learning for Mobile Robot Navigation from the Dynamical Systems Perspective. *IEEE Transaction on SMC (B), 26*(3), 421–436.
Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction process. *Neural Networks, 16*, 11–23.
Tani, J., & Fukumura, N. (1994). Learning Goal-directed Sensory-based Navigation of a Mobile Robot. *Neural Networks, 7*(3).
Tani, J., Ito, M., & Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB. *Neural Networks, 17*, 1273–1289.
Tani, J., Nishimoto, R., Namikawa, J., & Ito, M. (2008a). Codevelopmental learning between human and humanoid robot using a dynamic neural network model. *IEEE Transaction on System, Man and Cybernetics, 38*(1), 43–59.
Tani, J., Nishimoto, R., & Paine, R. (2008b). Achieving "organic compositionality" through self-organization: Reviews on brain-inspired robotics experiments. *Neural Networks, 21*, 584–603.
Tani, J., & Nolfi, S. (1999). Learning to perceive the world as articulated: An approach for hierarchical learning in sensory–motor systems. *Neural Networks, 12*, 1131–1141.
Vogt, S. (1995). On relations between perceiving, imaging and performing in the learning of cyclical movement sequences. *British journal of Psychology, 86*, 191–216.
Wolpert, D., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks, 11*, 1317–1329.
Yamashita, Y., & Tani, J. (2008). Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment. *PLoS Computational Biology, 4*(11).
Ziemke, T., Jirenhed, D., & Hesslow, G. (2005). Internal simulation of perception: Minimal neurorobotic model. *Neurocomputing, 68*, 85–104.