New Mathematics and Natural Computation, Vol.5, No.1, pp.307-334, 2009.

New Mathematics and Natural Computation

Creating novel goal-directed actions at criticality: A neuro-robotic experiment

Hiroaki Arie

RIKEN Brain Science Institute
2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan, arie@brain.riken.go.jp

Tetsuro Endo

Department of Advanced Mechanical Engineering, Waseda University
3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555 Japan,

Takafumi Arakaki

Department of Advanced Mechanical Engineering, Waseda University
3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555 Japan,

Shigeki Sugano

Department of Advanced Mechanical Engineering, Waseda University
3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555 Japan,

Jun Tani

RIKEN Brain Science Institute
2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan, tani@brain.riken.go.jp

Received 15 8 2008
Revised 11 9 2008

The present study examines the possible roles of cortical chaos in generating novel actions for achieving specified goals. The proposed neural network model consists of a sensory-forward model responsible for parietal lobe functions, a chaotic network model for premotor functions and prefrontal cortex model responsible for manipulating the initial state of the chaotic network. Experiments using humanoid robot were performed with the model and showed that the action plans for satisfying specific novel goals can be generated by diversely modulating and combining prior-learned behavioral patterns at critical dynamical states. Although this criticality resulted in fragile goal achievements in the physical environment of the robot, the reinforcement of the successful trials was able to provide a substantial gain with respect to the robustness. The discussion leads to the hypothesis that the consolidation of numerous sensory-motor experiences into the memory, meditating diverse imagery in the memory by cortical chaos, and repeated enaction and reinforcement of newly generated effective trials are indispensable for realizing an open-ended development of cognitive behaviors.

Keywords: novel goal-directed action; chaotic dynamics; criticality; CTRNN.

1. Introduction

Why some cognitive acts entail explicit consciousness but many others undergo unconsciously? For example, we can grasp a coffee mug without paying much attention to the action, while consciously talking with others. Some neuroscience researchers have considered that consciousness is deeply involved with prefrontal lobe activity. It is well known that prefrontal lesion patients often have problems in generating new plans to achieve novel given goals even though although they seemingly have no problems in generating everyday's skilled actions such as reaching for a mug and grasping it[1,2,3]. The former case involves the conscious and deliberate manipulation of mental images in planning the novel goal-directed actions which presumably require prefrontal activity. On the other hand, the latter case seems to proceed automatically with having less activities in the prefrontal lobe. How can we model the underlying mechanisms accounting for these phenomena? The conventional cognitive science scheme might not be appropriate here since its explicit computational scheme is too formal to describe implicit tacit knowledge for the generation of unconscious skilled behaviors as well as pragmatic aspects in creating novelty in ideas and actions in human and animals.

The current paper attempts to look at this problem by using an alternative scheme involving a dynamical system approach[4,5,6,7] combined with robotics synthetic experiments. In particular, we focus on the possible roles of cortical chaos in generating "creative" behaviors as has been inspired by the pioneering studies of Walter Freeman who discovered chaos in memory dynamics in the olfactory bulb in animals[8]. Our group has conducted neuro-robotics researches[6,9] since most cognitive phenomena might be better understood by considering the coupling between the internal neuronal dynamics and that one of the body and environment.[10]. We have shown that certain compositional structures can be self-organized in the internal neuro-dynamic memories as regularities hidden in the sensory-motor interactive experience are consolidated[6,11]. The current paper will attempt to extend this line of thought to the problem of how novel action imagery[12] can be "created" from the memory dynamics of consolidated.

Firstly, we provide a review of our ideas[11] regarding how multiple goal-directed actions can be learned as skilled ones through repeated practice and then generated by reviewing our neuro-robotics studies on the tasks of object manipulation using vision. We have considered that the trajectories of different goal-directed behaviors can be generated depending on the initial states given to a particular neural dynamical system by means of the initial sensitivity characteristics of the nonlinear systems[13].

The neuro-scientific interpretation of this hypothesis is that the ventral premotor (PMv), which is regarded as a dynamical system, is set with its initial state corresponding to the currently specified goal state. Then, the PMv dynamics generates a corresponding abstract action plan in terms of temporal sequences of motor acts imagery. The generated image of sequences of motor acts are fed into the inferior parietal lobe (IPL) after which the IPL generates a look-ahead prediction of the corresponding visuo-proprioceptive (VP) flow by means of the so-called "sensory-forward model"[11,13]. This scenario can be rephrased as follows: PMv generates an abstract scenario for achieving a given goal while IPL anticipates the detailed sensation based on the scenario. It should be noted that our assumption of the anticipation function for IPL might be unusual to some readers since the main role of IPL has been regarded as an integrator of different modalities of sensory input[14]. However, there has been some evidence[15,16,17,18] supporting our ideas as described in detaile in the next section. Our assumption for the role of PMv accords with the main arguments proposed by the Rizzolatti group[19] that the mirror neurons encode the underlying goals of movements rather than exact movement profiles.

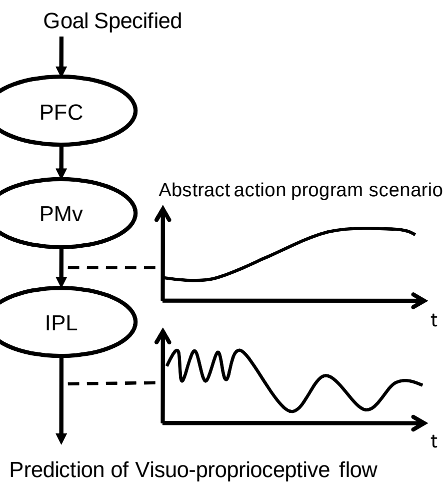


Fig. 1. The top-down pathway to predict visuo-proprioceptive flow in the proposed three-leveled architecture

Our model architecture consists of three-level dynamical systems as shown in Figure1. First, the prefrontal lobe initiates the PMv forward dynamics by setting its initial state after which the PMv generates an abstract action scenario as a slowly changing temporal pattern. Then, by receiving these patterns as input from the PMv, the IPL anticipates the exact VP profile by utilizing its forward dynamics of having faster time-constant.

In our robotics setting, with a given action program in terms of slowly changing patterns, the sensory-forward model generates a forward prediction of the next VP sensory state from the current one, where the visual state represents the visual image of the currently attended object, and the proprioceptive state represents the posture of the arm. Therefore, for a given goal, such as grasping a mug, the sensory-forward model can generate a predicted image of how the arm postures as well as visual image of a mug changes in the course of grasping the object. Then, the predicted arm movement image can be enacted by inversely computing the necessary motor torque in the cerebellum and the motor cortex. Here, the entire process of generating a motor behavior from a given goal can be carried out automatically without any deliberation in the sense that certain temporal patterns are developed spontaneously in combined neural network dynamics from a given initial state.

Next, we consider how brains can generate novel combinations of action programs for achieving novel goals. One possibility is to utilize stochastic noise to combine parts of learned trajectories into novel ones. The current paper, however, investigates an alternative possibility, in which deterministic chaos plays an essential role in generating novel combinations. If a forward model is learned by a set of target sequences utilizing the initial sensitivity, the mapping of initial states to generated temporal patterns would be simple, as has been shown in Ref.[13]. However, if the forward model is learned as being associated with certain networks whose dynamics is characterized by chaos, the mapping can become complex by means of the strong initial sensitivity of the chaos. In such situations, the forward model can generate diverse imagery of novel sequences by combining and modulating the prior-learned temporal patterns depending on the initial state. Then, it might be possible to search for the initial state, which leads to the generation of novel action programs satisfying newly given goals. This search process of the initial state by chaos might correspond to the deliberation of action planning which has been assumed to take place in the prefrontal cortex. In particular, if the search process encounters difficulties with respect to finding the best match for the goals, this would entail consciousness. It is intuitive that truly novel or "creative" ideas are likely to appear with certain criticality in our everyday lives. Also, we often experience that many of such novel ideas could be ineffective in reality. Therefore, we know that tentatively generated novel plans or images have to be examined by enacting them in the reality. This might be the same even for our simple experimental robots who seek novel action plans.

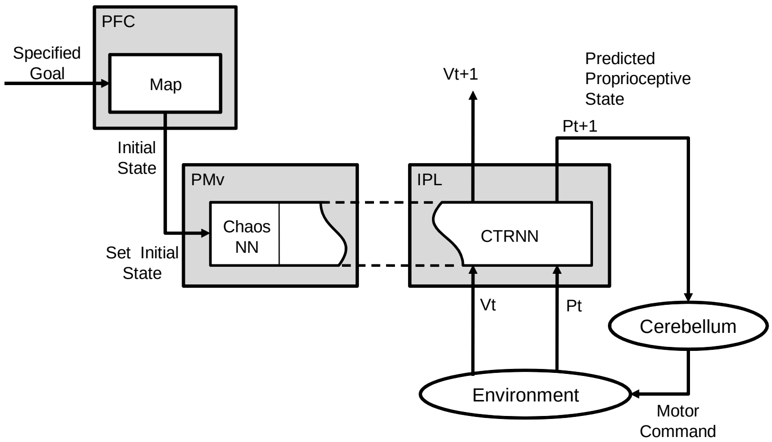
The present paper describes neuro-robotics synthetic experiments for embodying the above-mentioned ideas. A small humanoid robot is trained for a set of different goal-directed actions by tutors. First, after consolidating the trained experiences to the internal memory, we examine how trained actions can be regenerated automatically. Then, we examine our main focus of how the robot can generate diverse imagery of possible plans to satisfy given novel goals by utilizing prior-acquired memory and how they can be enacted in the reality. The next section begins with a description of the proposed model.

2. Model

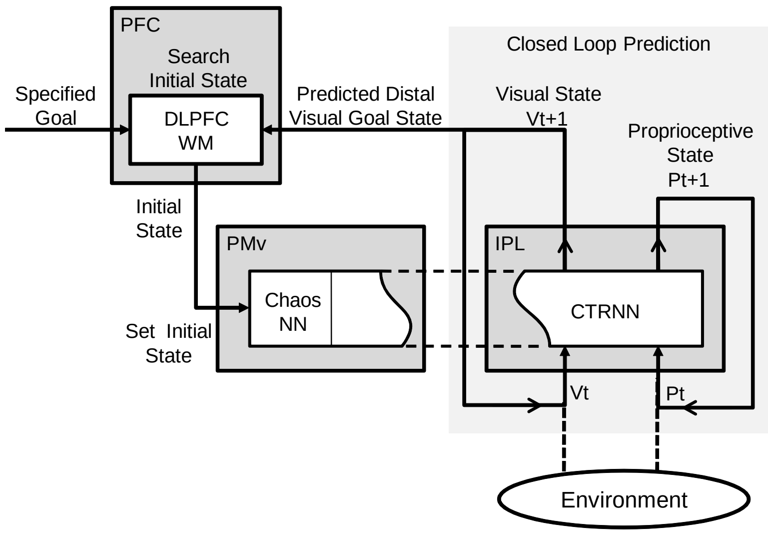
This section introduces the proposed model with accounts for its biological plausibility, the neural network mathematics and the implementation details of humanoid robot.

2.1. *Brain Model*

Figure2 shows the overall schematics of how each part of the brain functions and interacts with other parts in the proposed model.



(a) Execution of skilled goal-directed actions



(b) Searching for a novel goal-directed action plan

Fig. 2. Correspondences of the model to the brain anatomy where the execution of skilled goal-directed actions in (a) and searching for action plans for novel goal states in (b)

2.1.1. *Generation of well-trained actions*

Firstly, let's look at the case of generating well-trained goal-directed actions, as shown in Figure2 (a). When a goal to be achieved is specified from well-trained ones, its corresponding initial state is set to PMv by remembering it. Both of the PMv and IPL are implemented by a continuous time recurrent neural network (CTRNN) model[20,21]. The CTRNN dynamics in the PMv is designed to be chaotic with pre-wired synaptic connectivity. The synaptic weights inside the chaotic network are designed such that the maximum Lyapunov exponent becomes positive, which will be detailed later. On the other hand, the synaptic connectivity in the IPL is determined through the learning processes.

By having connectivity between these two networks, the IPL learns the details of the VP patterns by predicting the VP state at the next step (V_{t+1}, P_{t+1}) from the previous one (V_t, P_t), while the PMv tends to combine and modulate those patterns memorized by IPL through interacting with it. The central assumption in the current model is that brains might utilize cortical chaos in enhancing diversity of the behavioral patterns generated by utilizing its initial sensitivity characteristics. The initial state in terms of the neural activation states in the chaotic neural network is manipulated in order to generate different goal-directed actions.

The training of IPL is conducted in a supervised manner by utilizing the VP sequences acquired as input to the IPL through the tutor guidance of the robot. The error signal is back-propagated among IPL to PMv by which the synaptic weights in the IPL network are tuned. It is important to note that the learning proceeds with accompanying the pre-wired chaotic dynamics such that different initial states of the chaotic network initiate trajectories of different trained goal-directed actions. In our proposed learning scheme these initial states as well as the synaptic weights in IPL are self-determined in the course of training with a set of goal-directed action trajectories. The time constant of the chaos dynamics in PMv is set relatively to be larger than the one in IPL such that two levels of functions dealing with abstract plan scenarios and primitive behavioral pattern details can emerge. In the recent study of our group Yamashita and Tani[22] showed that such functional decomposition can emerge in the neural network model of multiple time scales RNN (MTRNN).

2.1.2. *Generation of novel actions*

Next, we describe the case of generating novel goal-directed actions. As we discussed in the previous sections, some goal-directed behaviors which have not yet been experienced can be generated by combining partially learned motor acts. This might be conducted by searching of the initial state and examining the imaginary generation of VP trajectories. Let's see the ideas by looking at Figure2 (b). Firstly, the goal specified externally is stored in the working memory assumed in the dorsolateral prefrontal cortex (DLPFC). Then the distal goal image which is generated by the forward dynamics with a particular initial state is matched with the specified goal image in the working memory. This is performed in the so-called closed-loop mode without the actual execution of motor acts in which the consequences of one's own actions are mentally simulated[6,9,23] . The forward dynamics is conducted without the actual sensory input but with the sensory imaginary loop utilizing its own prediction as shown by the dotted lines in Figure2 (b). If matching the specified goal image and the generated one fails, another initial state is examined. This search continues until perfect match is achieved.

If perfect match is obtained with a certain initial state, the actual movement is initiated by setting this initial state to PMv in the same way as described in Figure2 (a). It is also possible to perform on-line planning of a given goal-directed actions after actual movements are initiated. In this case the current activation states of the same neural units utilized for the initial state setting is manipulated for searching for the best match with the goal in an on-line manner. This sort of on-line planning becomes important when the environmental situations tend to change dynamically.

2.2. *Mathematical Modeling*

In the following sections, we describe the mathematical details of MTRNN with chaotic dynamics network model. The MTRNN model used in this study is shown in Figure3. We used a small humanoid robot which interacts with the environment with utilizing vision. The neural network model receives the proprioceptive state representing the posture of the robotic arm in terms of the joint angles, and visual state representing the direction of the camera and the image obtained with the camera. At first, these sensory inputs are pre-processed using topology preserving maps (TPM), which transform a vector of continuous values into neural population coding. This type of representation is considered to correspond to the neural representation in primary sensory cortices, such as V1 and S1.

After this transformation, input signals encoded by population coding are sent to the MTRNN. The role of the MTRNN is to predict the VP state at the next time step on the basis of the current one. This prediction is made possible by the capability of the MTRNN to preserve the internal state associated with complex dynamics. The actual robot arm movement just follows this prediction of the proprioceptive state.

In our current model, the MTRNN consists of three parts, namely input-output network, context unit network and chaos network. Both the input-output network and the context unit network correspond to the IPL. In addition, the chaos network corresponds to the PMv, which plays an important role in generating goal directed actions, as describes in previous section.

Since chaos has strong initial sensitivity characteristics, diversity of temporal sequences can be generated by setting different initial state values for the chaos network. On the other hand, if exactly the same initiail state is given to the chaos network, it always generates the corresponding temporal sequences. The training of multiple goal-directed actions is conducted utilizing this initial sensitivity and

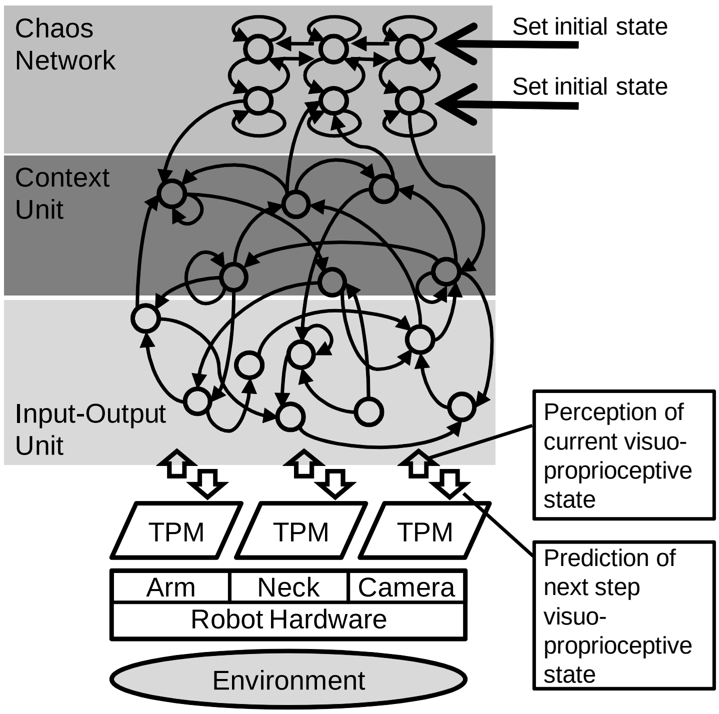


Fig. 3. Neural network model interfacing with a robot

deterministic characteristics. It is worth mentioning that the chaos neural network is easily perturbed by the external noise in the environment when generating actions using physical robot in our current study.

The teaching VP sequence is obtained through the manual guidance of the robot arms for each target goal-directed action. Each VP sequence sampled through the guidance is utilized for adapting the connective weights in the TPMs, the input-output network and the context unit network. In addition to adjusting the connective weights, the optimal initial state value in the chaos network is determined for reconstructing each VP sequence. After the successful training of the networks, the goal-directed actions can be regenerated by setting the specific initial state values.

2.2.1. *Implementation in a humanoid robot*

The robot has a head, which is equipped with a stereo camera, and two arms, each of which has 4 rotational joints. Our neural model receives the following sensory-input from the robot, namely the proprioception P_t (an 8-dimensional vector representing the angles of the arm joints), the direction of the camera on the head V_t^d (a 2-dimensional vector representing the rotational angle of the neck joints), and the visual perception V_t^v (16×12 dimensional vector representing the retinal image).

Each pixel component of this 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 the corresponding retinal cell, respectively. The direction of the camera is controlled by a PID controller which is programmed to track a red-colored object to be centered in the retina image. Receiving the current VP state (P_t, V_t^d, V_t^v), the neural network generates the next step prediction of them as $P_{t+1}, V_{t+1}^d, V_{t+1}^v$. Then, the robot arm is moved toward P_{t+1} as target joint angles with a PID controller.

The sensory-inputs of the proprioception, the direction of the eyes and the visual perception are initially processed by the corresponding TPMs. The TPM has characteristic that the topological properties of the input space is transformed into the sparsely activated population map[24,25]. Therefore this sparse encoding transformation reduces the overlap of the VP trajectories. The size of the TPMs is 144 (12×12) for proprioception, 36 (6×6) for the direction of the camera head and 100 (10×10) for the visual perception. The mathematical details are described in the paper written by Yamashita and Tani[22].

2.2.2. *Forward generation of MTRNN*

We designed our neural network model using the multiple time constant recurrent neural network (MTRNN) model[22]. The MTRNN is a type of the CTRNN model[20,21], in which neurons have multiple time constants. This difference with respect to time constant plays an important role in self-organizing the functional hierarchy structure. Essentially, the MTRNN takes input as different modalities of sensation and mingles those inputs together to generate predictions of their time development in the future. This model has the ability to segment VP sequences into reusable primitives and self-organize a functional hierarchy of them while learning complex VP sequences. When generating an action, these reusable motor primitives are flexibly integrated into various patterns of complex VP sequences. The effect of having multiple time constants among neurons is roughly summarized as follows: neurons with fast time constant remember each primitive and neurons with slow time constant switch between these primitives.

The MTRNN has three groups of neural units in our current model, namely input-output units (280), context units (80) and chaos network units(20). The number inside each parenthesis indicates the number of units for each network. Among the input units, the first 144 units ($i = 1 - 144$) correspond to proprioceptive input P, the next 36 units ($i = 145 - 180$) correspond to the direction of the camera (V^v) and the rest ($i = 181 - 280$) correspond to visual perception (V^d). In the following section, r denotes this 280-dimensional vector transformed by TPMs using the actual state of the robot.

Synaptic weights within the chaos network are pre-defined and fixed. These neurons in the chaos network receive input signals from context neurons. The neural units in the chaos network and the context units are mutually connected with randomly determined fixed synaptic values. The context units are connected with the input units of which synaptic weights are determined through learning. The activation of these neurons is calculated using the following equations.

$$\tau_i \frac{du_{i,t}}{dt} = -u_{i,t} + \sum_j w_{ij} x_{j,t} \tag{2.1}$$

where $u_{i,t}$ is the potential of each i-th neural unit at time step t and $x_{j,t}$ is defined as follows

$$x_{j,t} = \begin{cases} r_{j,t} & \text{if } i \in O \\ y_{i,t-1} & \text{otherwise} \end{cases} \tag{2.2}$$

At the i-th neural unit, the signal from the j-th presynaptic neural unit has a weight of w_{ij}. In Eq.2.2, $y_{i,t}$ represents a neuronal activation of the i-th neural unit at time step t, and $\tau$is the time constant of the neural unit. This time constant affects the response rate of the neuronal activation. If the parameter τ is small, the potential of the neural unit can change rapidly. Otherwise, the change is slow. The time constants are set to 2.0 and 3.0 for input-output units and context units, respectively. The neural activation for i-th neuron, $y_{i,t}$, is calculated using the following equation

$$y_{i,t} = \begin{cases} \frac{\exp(u_{i,t})}{\sum_{j \in Z} \exp(u_{j,t})} & \text{if } i \in Z \\ f(u_{i,t}) & \text{otherwise} \end{cases} \tag{2.3}$$

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.4}$$

where Z is P or V^d of V^v. The update rule of the internal activation state, $u_{i,t}$, for each integration time step is given by discretizing Eq.2.1 where the time step interval Δt is taken as 1.

$$u_{i,t+1} = (1 - \frac{1}{\tau_i}) u_{i,t} + \frac{1}{\tau_i} \sum_j w_{ij} x_{j,t} \tag{2.5}$$

The neuronal activation of context neurons is calculated by using the conventional sigmoidal function. Alternatively the input-output neurons are calculated using softmax activation, which allows the MTRNN to maintain consistency with the TPM output. The output vector of the MTRNN is sent to the TPM and subsequently transformed into the predictions of proprioception P_{t+1}, the direction of the eyes V^d_{t+1} and the visual perception V^v_{t+1} using Eq.**??**. After this, the robot initiates movement using these predicted values.

2.2.3. *Learning of synaptic weights*

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 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,

$$E = \sum_{t} \sum_{i \in O} y^*_{i,t} \log\left(\frac{y^*_{i,t}}{y_{i,t}}\right) \tag{2.6}$$

where $y^*_{i,t}$ 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 weight. As a training method, we use the general Back Propagation Through Time (BPTT) algorithm[26]. 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 $\partial E/\partial w$. In the actual learning process, the update rule of a connective weight from the i-th neuron to the j-th neuron at the n-th learning iteration step is as follows,

$$w_{ij}(n+1) = w_{ij}(n) - \alpha \frac{\partial E}{\partial w} \tag{2.7}$$

where α is the constant parameter, which adjusts the learning rate. The gradient $\partial E/\partial w$ is given by following equations.

$$\frac{\partial E}{\partial w} = \sum_{t} \frac{1}{\tau_i} \frac{\partial E}{\partial u_{i,t}} y_{j,t-1} \tag{2.8}$$

$$\frac{\partial E}{\partial u_{i,t}} = \begin{cases} y_{i,t} - y^*_{i,t} + (1 - \frac{1}{\tau_i}) \frac{\partial E}{\partial u_{i,t+1}} & \text{if } i \in O \\ \displaystyle\sum_{k \in N} \frac{\partial E}{\partial u_{i,t+1}} \left[\delta_{ik}(1 - \frac{1}{\tau_i}) + \frac{1}{\tau_k} w_{ki} f'(u_{i,t})\right] & \text{if } i \notin O \end{cases} \tag{2.9}$$

where $f'()$ is the derivative of the sigmoidal function and $\delta_{ik}$is Kronecker's delta ($\delta_{ik} = 1$ if $i = k$, otherwise $\delta_{ik} = 0$).

Through the iterative calculation of the BPTT, the values of the connective weights reach to their optimal values in the sense that the error between the teaching sequences and the output sequences is minimized. Throughout the learning iteration, the learning rate α is fixed at 0.0008. The initial values of the connective weights were set with random values ranging from -0.1 to 0.1

In the BPTT calculation, the predicted value of P_t, V^d_t, V^v_t can serve as imaginary sensory feedback for the next time step $t+1$ instead of the actual feedback from the robot movement. We call this imaginary sensory feedback as mental simulation, as described in the previous section. During the training iterations, this mental simulation is also mixed with actual sensory feedback in order to accelerate the convergence. Specifically, we use the following equation.

$$\hat{P}_{i,t+1} = 0.9P_{i,t+1} + 0.1P^*_{i,t+1} \tag{2.10}$$

$$\hat{V}^d_{i,t+1} = 0.9V^d_{i,t+1} + 0.1V^{d^*}_{i,t+1} \tag{2.11}$$

$$\hat{V}^v_{i,t+1} = 0.9V^v_{i,t+1} + 0.1V^{v^*}_{i,t+1} \tag{2.12}$$

where $\hat{P}_{t+1}, \hat{V}^d_{i,t+1}, \hat{V}^v_{i,t+1}$ represents the virtual sensory feedback for the next timestep, $P_{i,t+1}, V^d_{i,t+1}, V^v_{i,t+1}$ represents the predicted value and $P^*_{i,t+1}, V^{d^*}_{i,t+1}, V^{v^*}_{i,t+1}$ represents the actual sensory feedback, respectively.

For both learning and generation, the initial states of the context units, except for two units in the chaos network, are set to their neutral values, i.e., the potential of those units are set to 0. On the other hand, the initial activation states of the two chaos units are self-organized during the learning iteration. The details of this method are described in the next section.

2.2.4. *Chaos Network*

As mentioned previously, our neural model has pre-defined connective weights in the chaotic network. These connective weights are defined in the way dynamic activities of these neurons should be chaotic, meaning that the maximum Lyapunov exponent becomes positive. One way to implement a chaotic neural network is based on Ruelle-Takens-Newhouse scenario that is to connect multiple neural oscillators having different time constants. As shown in Figure4, the neural oscillators, which

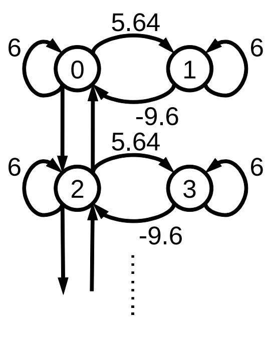


Fig. 4. Chaotic oscillator

consisted of two neurons, are connected through neurons which have even index numbers. In the oscillator, each two neurons which have the same time constant are connected with connective weights of 5.64 and -9.6, as shown in Figure4. The neuronal activation of these neurons is calculated using the following equation.

$$y_{i,t} = f(u_{i,t} + \theta_I) \tag{2.13}$$

$$\frac{du_{i,t}}{dt} = -u_{i,t} + \sum_j w_{ij} x_j \tag{2.14}$$

where $f()$ is the sigmoidal function defined in Eq.2.4 and x_j is neuronal activation including input-output and context neurons. The bias values θ_i are set as follows

$$\theta_i = \begin{cases} 0 & \text{if } i \text{ is even} \\ -6 & \text{if } i \text{ is odd} \end{cases} \tag{2.15}$$

In our current model, there are 20 neurons in the chaotic neural network, and the initial activational states of 18 neurons are set as neutral for both training and generation. The initial activational state for the other two neurons is self-organized during the learning iteration. The learning rule for these two initial values is based on the gradient descent method. In the learning phase, the squared error, defined as following equation, is evaluated for each target behavior.

$$E_s = \sum_t \sum_{i \in O} (y_{i,t}^{k*} - y_{i,t}^k)^2 \tag{2.16}$$

where $y_{i,t}^{k*}$ is the desired activation value of the output neurons represented by population coding, and $y_{i,t}^k$ is the activation value of the output neurons with current connective weights. In order to calculate the local gradient information, small perturbation is added to the initial activation value of two neurons in chaos network. Then, the initial activation value of the two neurons is slightly shifted in the direction of the negative gradient at that point.

2.2.5. *Planning by the initial state search*

In our current study, the goal state is specified by the corresponding visual state (V_{goal}^v, V_{goal}^d). The neural network generates multiple imaginary VP sequences by closed-loop operation with sweeping the 2-dimensional initial state space. Among these multiple VP sequences, the Euclidean distances between the distal visual state of the imaginary VP sequence $(V_{distal}^v, V_{distal}^d)$ and the specified visual state (V_{goal}^v, V_{goal}^d) are calculated. Here, the distal state is defined as a state at a particular time step where the Euclidean distance takes a minimum value within the predefined step length. The distance is calculated using vectors represented as a population coding and the best match is taken by finding the closest VP sequences.

3. Experiment

3.1. *Task design*

The experiment was carried out by using a small-scale humanoid robot named HOAP3. The robot was fixed to a chair, and a workbench base 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 base immediately in front of the robot. Also, a low height pedestal was fixed to the base behind the object. The robot was required to learn a set of goal-directed actions of manipulating the object. A human tutor prepared teaching target VP sequences used for training of the neural network model. In these goal-directed behaviors, the robot started to move from the same starting posture while the object was placed as either standing or lying on the base in immediately in front of the robot.

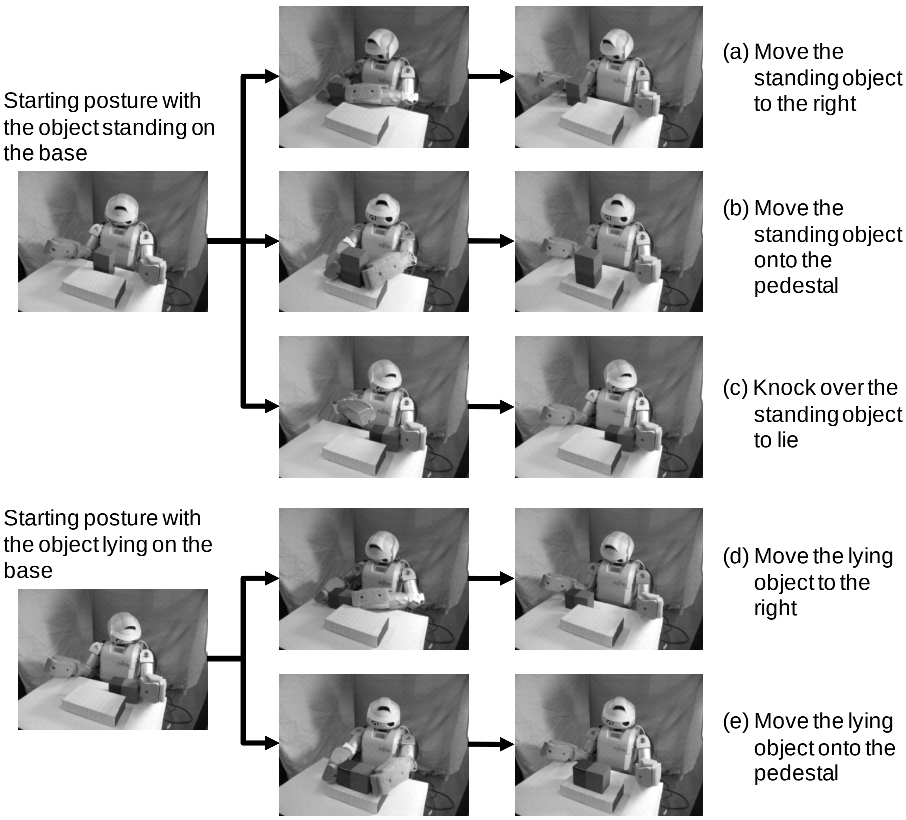


Fig. 5. Five types of goal-directed actions

The robot was tutored for five operational actions, which are shown in Figure5: (a) move the standing object to the left, (b) hold up the standing object and put it onto the pedestal, (c) knock over the standing object to lie, (d) move the lying object to the right, (e) hold up the lying object and put it onto the pedestal. The object was placed in the center in front of the robot in actions (a-c) and in the left-hand side in actions (e,f). For each operational action, the robot was tutored 3 times with changing the initial object position as 2cm left of the original position, the original one and 2cm right of the original one. Therefore, a total of 15 VP sequences were sampled for the training of the network.

After the training, two types of the robot behavior generations were examined. The first one was simply regenerate the 5 trained goal-directed actions. The network was set with the corresponding initial state values which had been self-determined for each goal state with the case of the original object position in the training phase and then the robot was started to move. The robot movement was tested for the left, the original and the right object position cases for each goal-directed action.

The second one is to plan and to generate a novel goal-directed action. The goal state is given as the visual state of the object lying on the pedestal with the condition that the object is initially standing on the base in front of the robot. Note that this combination of the goal state and the object initial position is novel for the robot, i.e., the robot has to generate novel motor-act sequences of firstly knocking down the object to lye on the base and then holding up this onto the pedestal. In this experiment, action plans are generated by searching for the initial state which can bring the best match between the specified goal state and the predicted one in the distal step by the sensory-forward model. The search is iterated 100 times with a 0.1 grid in the two-dimensional initial state space in the range between 0.05 and 0.95 for each dimension. Then, the plan of the best match is enacted by the robot in the physical environment with activating the network forward dynamics with the corresponding initial state values.

3.2. *Results*

3.2.1. *Regeneration of the trained goal-directed actions*

All three Kohonen networks were pre-adapted before the training of the MTRNN utilizing the sensory patterns acquired for the 15 tutoring sequences. For the training of MTRNN, BPTT was iterated for 10000 epochs with randomly set initial weights. The mean square error (MSE), which is the average square error per input-output neurons per step over all teaching VP sequences, converged to 0.000179.

The networks after the training were tested to regenerate the trained actions by the robot. Each goal-directed action is tested with varying the object position among the left, the original, and the right ones. Each trial was repeated for three times. The results of the experiment show that the robot could regenerate 15 actions (3 positions $\times$ 5 actions) with a success rate of about 70%. The details are shown in the Table1. As shown in this table, the success rate was 93% when the object was

Table 1. The number of successful trials within three trials in regeneration of the trained actions.

Specified goal-directed action	Position of the object		
	2cm left	original	2cm right
(a) standing object to right	3	2	2
(b) standing object onto the pedestal	0	3	1
(c) standing object to lying	3	3	3
(d) lying object to right	0	3	2
(e) lying object onto the pedestal	3	3	0

placed in the original position. The score becomes lower when the object is located in the left or right positions. It can be said that in most cases the trained actions can be regenerated successfully. We found that the instability of light reflection on the colored object in the visual sensation caused certain fluctuations in the MTRNN dynamics during the task execution.

Figure6 shows the time developments of the goal-directed actions for the first three of the five representative cases in which the object is placed in the original position. The upper graphs show the activation of some neurons in the chaos network.

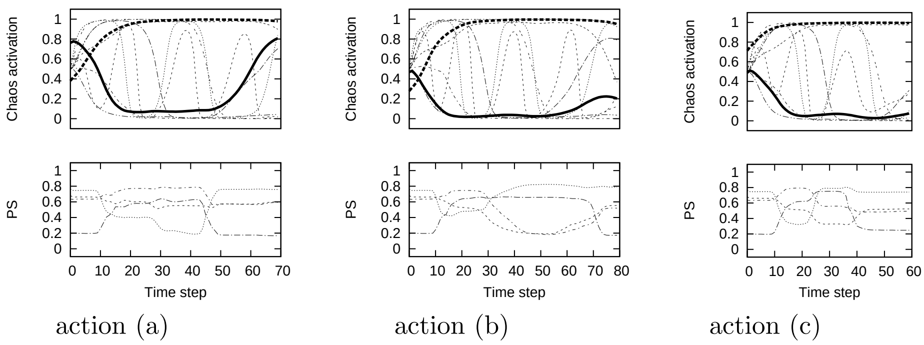


Fig. 6. Time developments of three goal directed actions. Each pair of three graphs corresponding to the three goal-directed actions (a-c) as described in Figure5. The upper graphs show the activation profile of the chaos network units. The lower graphs show the proprioceptive state (PS) for the four joint angles in the right arm.

Especially, the two bold lines (solid and a dotted) represent the activation of the neural units whose initial values were self-determined. The activation of these two neural units starts from different initial values depending on the goal. The lower 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.

3.2.2. *Planning and generating actions for achieving the specified novel goal state*

In this experiment, imaginary VP sequences were generated utilizing closed-loop operation, in which predicted sensory values serve as virtual sensory feedback, so the network can generate imaginary VP sequences without generating physical movements of the robot. As a result of searching the best match VP imaginary sequences with varying the initial state for 100 times, only one VP sequence was found to satisfy the specified goal state in its distal step image. The initial state of the successful sequence was (0.35, 0.95). We constructed the initial state map which shows the mapping from the two-dimensional initial state to the resulting VP imaginary sequences of 230 steps period. This map indicates how the mental imaginary sequences vary as sensitive to the initial state. The result can be seen in Figure7, where each grid is labeled in accordance to the generated behavioral pattern categories, where from "A" to "E" denoting the trained ones, "X" denotes an unclassifiable distorted one, and "Z" denotes frozen one i.e., the dynamics is stopped at equilibrium points. It is seen that most of these VP sequences appear to combine the trained ones as

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	DZ	EA	DZ	EX	BB	BA	BB	AB	XX	XAB	
0.1	DB	EC	EB	BB	BB	BB	AA	AC	XB	XB	
0.2	BX	DA	AC	BB	BB	BC	AA	AA	AA	EB	
0.3	BA	BB	BB	BB	BX	AA	AE	AA	BA	CB	
0.4	CA	CA	CA	CB	CC	CA	CA	CB	CA	CC	
0.5	CA	CB	CB	CB	CA	CA	CB	CA	CA	CA	
0.6	CA	CA	CA	CB	CA	CA	CB	CC	CA	CC	
0.7	CA	CA	CB	CB	CA	CA	CB	CA	CB	CA	
0.8	CA	CA	CB	CA	CB	CA	CA	CA	CB	CA	
0.9	CA	CA	CZ	○	CC	CA	CA	CB	CB	CA	
1.0											

Fig. 7. The initial state map obtained after the first training of the network. Each axis represents the activational values of two chaos neural unit. Each grid is labeled in accordance to the generated behavioral pattern categories where from "A" to "E" denoting the trained ones, "X" for unclassifiable distorted one, "Z" for frozen one and "○" denotes the best match VP sequence.

like "CA", "CB", "DB". It is observed that these combinations appear as clusters in several regions in the map. Among those "○" denotes the best match VP sequence imaged for the given novel goal state.

The upper part of figure8 shows the generated sequence of the visual imagery in the retina and the direction of the head camera represented as its corresponding TPM in the lower part of the figure, starting with the best match initial state (0.35, 0.95). It is seen that when the two-colored object is standing on the base at the 0th step, the object is knocked down by the right arm of the robot at the 30th step, the object is lying on the base with the left hand of the robot is in the air at the 45th step, the object is held up at the 90th step, then the object is put on the pedestal by both arms, and finally the object is lying on the pedestal after both arms release the object. It should be noted that the TPM population coding for the direction of the head camera is different between the two situations where the object is lying on the base at the 45th step and that where it is lying on the pedestal at the 210th step. Therefore, this imagery is regarded as a sequential combination of knocking over the object and then holding up the lying object onto the pedestal.

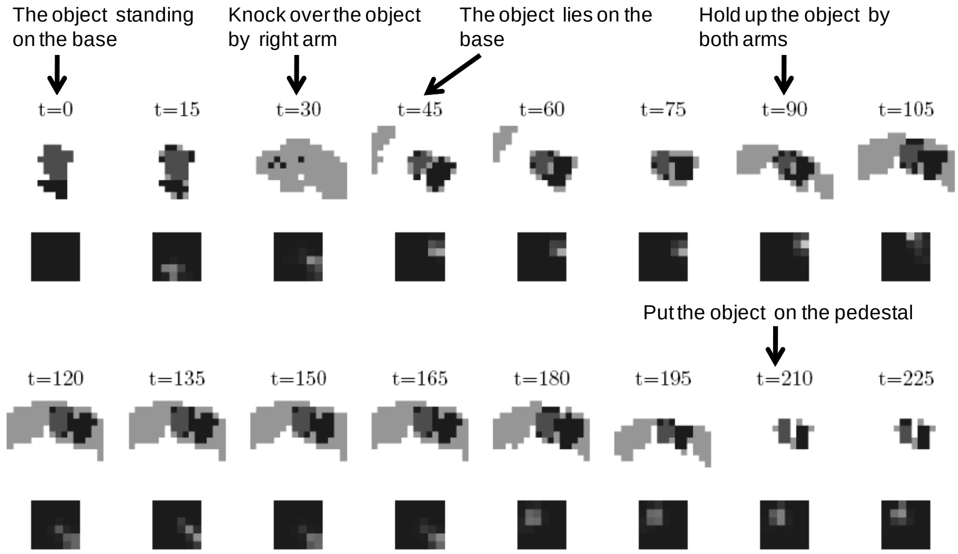
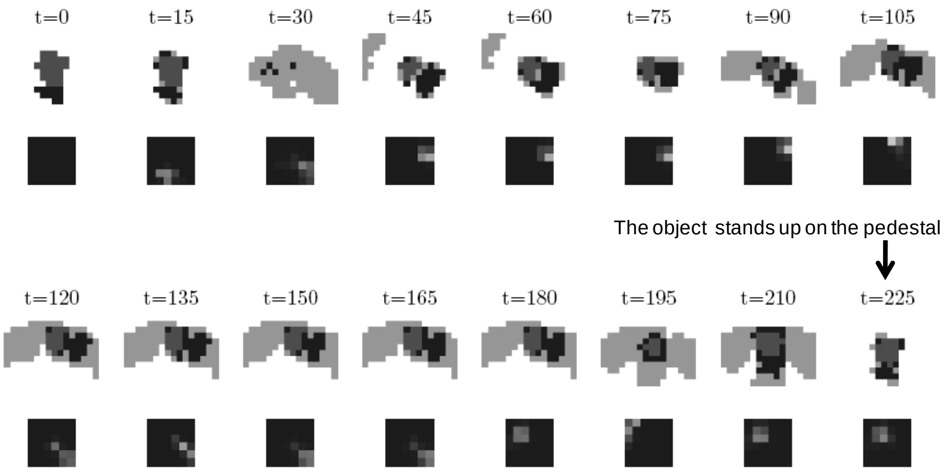


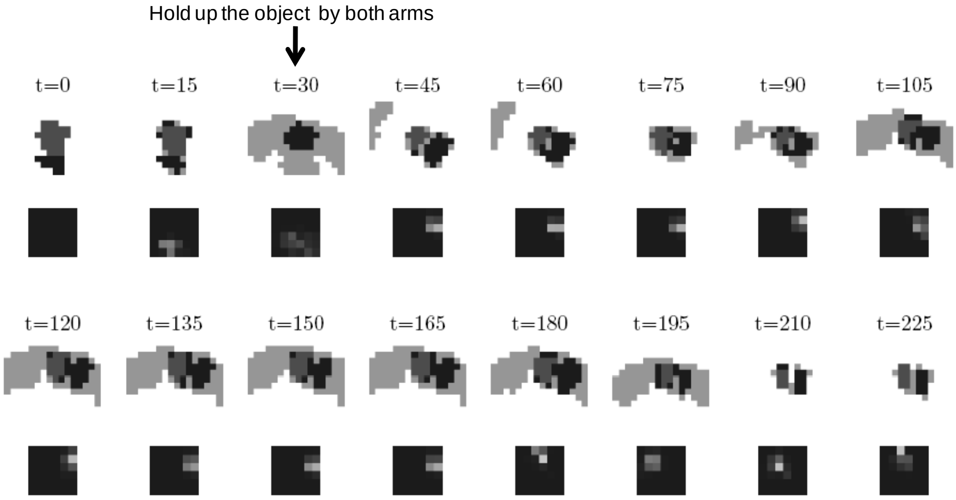
Fig. 8. Imaginary sequence generated with the initial state (0.35, 0.95) where the upper part shows the retinal image and the lower part does for the TPM population coding for the camera head direction.

However, it is true that this combination appears in only one grid: (0.35, 0.95). We further investigated the initial state sensitivity in generating the VP imaginary sequences within this grid in order to examine the stability of generating imaginary sequences of achieving the novel goal. It was found that even minor shifts of the initial state modulate the VP imaginary sequences substantially in this local region. Figure9 (a) and (b) show the imaginary sequences with the initial state set to (0.36, 0.96) and (0.32, 0.93), respectively. In the case of the initial state (0.36, 0.96) shown in Figure9 (a), it is observed that the imaginary sequence is mostly the same as the one with the initial state set to (0.35, 0.95). However, the object suddenly stands up from lying position at the very end of this imaginary sequence. In Figure9 (b), where the state is set to (0.32, 093), the generated VP imaginary sequence is completely different from the other two. In this case, the object was held and moved to the right-hand side of the robot. However, the object comes back to the original position by itself, so the robot repeated the same action. These results indicate that adequate VP imaginary sequences of achieving the specified novel goal state can be generated only by chance in the critical region in the initial state space where diverse imaginary sequences are generated. On the other hand, it was found that the generation of well-trained ones was much more stable without having keen initial state sensitivity around the regions determined in the training phase.

Next, we investigated whether the robot can physically achieve the specified goal state successfully by setting the initial state with (0.35, 0.95) of which plan



(a) Imaginary sequence with the initial state set to (0.36, 0.96)



(b) Imaginary sequence with the initial state set to (0.32, 0.93)

Fig. 9. Imaginary sequences generated with the initial state values within the vicinity of the best match case.

was found to be the best match. However, it was found that the robot could not achieve the goal with this initial state values. Next, the robot actions were generated repeatedly with varying the initial state within the found grid of (0.3-0.4, 0.9-1.0). It was found that successful accomplishment occurred in only two out of 40 trials. The VP sequence of one successful case is shown in Figure10. It was also found that the same initial state of one time successful case hardly repeats the same behavior

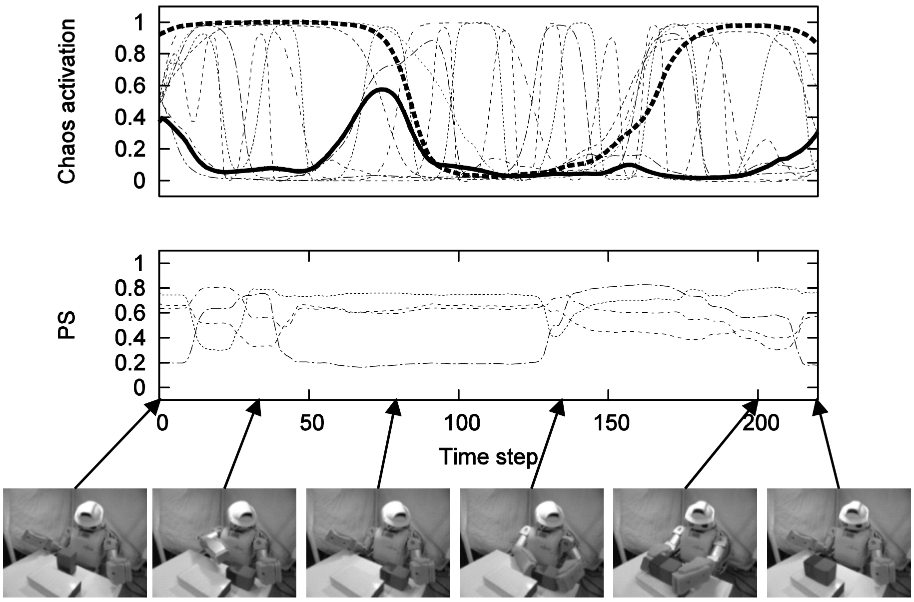


Fig. 10. Actual performance of the robot to achieve the given novel goal state. The upper graph shows the activation of chaos network units. The bottom graph shows the proprioceptive state (PS) for four joint angles in the right arm. Bottom pictures show the actual state of the robot at specific time steps.

patterns of the robot. It can be concluded that both the plans and their enactions for the given novel goal state are substantially sensitive to the initial state as well as to the external noise in the environment as compared to the cases of regenerating the well-trained ones.

3.2.3. *Additional reinforcement of effective actions generated*

Next, we examined whether the novel effective actions generated by chance in the previous section could be reinforced such that they could be regenerated in more stably. Two successful VP sequences which had been generated by the robot for achieving the specified novel goal were added to the original 15 VP sequences for additional incremental training. The additional training was conducted with 4000 BPTT iterations starting with the synaptic weights of the prior-trained one. The resultant MSE converged to 0.000167. Figure11 indicates the initial state map obtained after the additional training. It is seen that the regions corresponding to achieving the desired novel goal state (denoted by "◯" labels) are substantially enlarged in comparison to the one shown in Figure7. These regions, labeled as "C" in the previous map, are mostly converted to "◯". This is because action of "C", which means knocking over the object, is now reinforced to be followed by the ac-

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	EC	EC	XA	BA	BE	AA	XB	XE	XA	XA	
0.1	BA	BA	BE	BE	BE	AA	AA	AA	XE	XB	
0.2	O	CA	CA	BC	AA	BZ	AA	AX	AA	EA	
0.3	O	O	O	O	O	O	O	O	O	CA	
0.4	O	O	O	O	O	O	O	O	O	CA	
0.5	O	O	O	O	CA	CA	O	O	O	O	
0.6	O	O	O	O	O	O	O	CA	CZ	O	
0.7	O	O	O	O	O	O	O	O	CA	CB	
0.8	O	O	O	O	O	O	O	CA	O	CB	
0.9	O	O	O	O	O	O	O	O	O	CB	
1.0											

Fig. 11. Initial state map generated after the additional reinforcement of successful trials for the novel goal.

tion of holding up the lying object onto the pedestal. It was observed that mostly the same VP imaginary sequences are generated from these regions. When those plans were enacted by setting the initial state with those values, the robot was able to achieve the specified novel goal state with more than 70% success rate. This experimental result suggests that the reinforcement of successful trials by chance can increase the stability of regenerating them.

3.2.4. *Replacing the chaos network with a limit cycling network*

In order to clarify the functional role of the chaos network, we examined how the ability to generate novel action plans changes if the current chaos network is replaced with a networks which are characterized by limit cycling dynamics. The result indicated that, the network with limit cycling dymanics could not generate novel actions that achieve the given goal state. It was also observed that the diversity of generating VP sequences decreased significantly. These results might indicate that chaos is essential in searching novel action images by generating diverse action imaginaries.

4. Discussion

4.1. *Criticality in creating novel goal-directed actions and their reinforcement*

Our experiments showed that chaos plays an essential role in generating diverse mental images for actions by utilizing its initial sensitivity characteristics. It is, however, also true that chaos can also generate false images which do not correspond to the reality. In the current experiment, it was found that the case of achieving the novel goal is quite sensitive to the initial state and that even minor perturbations of the initial state could lead to the generation of false behaviors. Novel and meaningful actions are likely found at such criticality because it is the source of inherent diversity.

It means that our model generates diverse VP sequences; however, the actual generations of successful novel goal-directed actions are quite sensitive to noises and the initial states. Therefore, it is crucial that robots enact the generated images in reality and evaluate their consequences. And if some of such trials are found to be effective, such experiences should be reinforced by additional learning for its future utilization. It was observed that the novel goal-directed actions turned to be skilled ones with gaining their robustness after their reinforcement in our experiments. Finally, those actions can be generated automatically by remembering their initial states without deliberations.

4.2. *Origin of novelty*

It is also worthwhile considering that where the novelty comes from. The automatic generation of novel actions is essential for any exploratory-based learning systems. The standard reinforcement learning scheme[27] utilizes noise for generating novel motor patterns. The most likelihood actions for achieving the goals in the current step are mutated to novel ones by adding external noise. If the outcomes of these novel actions turn out to be effective, they are reinforced.

In the current model, the novelty is originated not from external noise but from the nonlinearity emerged in the network memory dynamics. Each memory of a goal-directed action is considered to be embedded in a distinct basin of attraction which can be accessed with the key of a specific initial state. Here, it is noted that each memory exists not as independent one but as a relational one among others because each neurons in the network participates in generating all memories. If the number of goal-directed actions to be memorized are increased, the relational structures among those memories become highly nonlinear as described as "attractor crowding" by Walter Freeman[28].

In this situation, the network mediated by the intrinsic chaotic dynamics tends to generate a diversity of false memories depending on the initial state. The resultant initial state map exhibits a rugged landscape in which the VP imaginary patterns vary abruptly even with minimum shifts of the initial state in some specific local regions while they do not in other local regions. Under this sort of highly nonlinear condition, truly novel and "creative" imaginary patterns might be generated only at such critical points because these are the edge of near break down of the relational memory structure consolidated. It can be said that both effects of attractor crowding and the initial sensitivity of the intrinsic chaos contribute to the generations of such local criticality. It is again reminded that the criticality affords diversity in generating novel patterns but with the costs of losing its stability and robustness.

4.3. *Related studies*

The current study is related to many other studies focusing on the roles of chaos in brains, cognition and behaviors. It is known well that Walter Freeman's group is the first to find chaos in brains. Based on EEG measuring of olfactory bulbs of rabbits, it was shown that chaos appears in the so-called "I don't know" states for the current smell[8]. On the other hand the neuronal dynamics is trapped into attractor of limit cycle when the smell is identified. This result suggests that chaos works as a catalyst for the networks to search flexibly identity of unfamiliar smells in the memory. This idea seems analogous to our interpretations of chaos shown in the current study. When the robot faces novel goals, the robot has to search for action plans utilizing chaos. On the other hand, familiar goals do not require the search by chaos.

Our ideas of organizing memories of behavior scheme in a distributed and relational manner shares with that by Walter Freeman discussed in his recent book[28]. If the memory is organized with the distributed representation rather than with the local one in the neuronal ensembles of either in the sensory-forward model or in olfactory bulbs, the nonlinearity of the network dynamics is enhanced. Such nonlinearity could afford "gestalt" in perceptions as well as in generation of motor behaviors. Furthermore, novel and "creative" actional images should emerge from such "gestalt" as have been discussed in the previous section.

Tsuda and his colleague[29,30] have speculated that cortical chaos may exhibit so-called the "chaos itinerancy"(CI). In their proposed chaos neural network model, the memory search dynamics tends to visit one quasi-stable attractor to another with preserving long temporal correlations in its itinerancy. Our preliminary study using the current model network showed that the similar phenomena take place when the network is trained with a set of behavior primitives of cyclic patterns such as holding up and down, moving left and right and pushing and pulling an object. Also in our current study, most of imaginary VP sequences generated by our model are combination of traind actions. So we considerd that with starting from the different initial states, the dynamics travels among primitives embedded in different limit cycle attractors with specific orders. Because the dynamics characteristics of CI seem to depend largely on the nonlinearity of the pre-wired chaos network, it is interesting to explore the details of such relations in future study.

Kuniyoshi et al.[31] recently showed that a simple coupling between a group of central pattern generators (CPGs) and robot body dynamics tends to generate diverse and meaningful behavior patterns when the intrinsic dynamics of the CPGs is chaotic. Although the proposed model by Kuniyoshi is still in a primitive form without having any internal models or mental simulation capability, their results have inspired the current study.

4.4. *Future studies*

Future studies should focus further more on dynamics of the incremental learning for novel goal-directed actions. The current study showed the case of only one-time reinforcement learning. It is, however, not clear that how much this learning can be iterated for acquiring further more novel actions. There should be a limitation for the memory capacity for this sort of the incremental learning. If so, do they scale with number of neurons allocated in the proposed model? However, more interesting question would be how the criticality is developed in the course of the incremental learning. It might happen that novel action can be recursively generated as a bootstrap from the previous learning of other novel actions with creating another critical region in the initial state map. However, in the course of packing new experiences by the bootstrap, do we expect to have a catastrophic breakdown of the whole memory system as like the self-organized criticality phenomena observed in the sand pile experiments by Bak[32]. If so, how does the system behave immediately before the catastrophic breakdown and after? This question should be related to understanding the basic principle of open-ended development of both biological and artificial cognitive systems and left for future studies.

5. Summary

The current study proposed a synthetic model of representing how well-skilled actions as well as novel ones can be generated in goal-directed ways by utilizing neuro-dynamical systems characteristics. The proposed hypothetical model consists of the sensory-forward model assumed in IPL, a chaotic network model assumed in PMv and prefrontal cortex model responsible for manipulating the initial state of the chaotic network.

Our experiments using a small humanoid robot implemented with the model showed that (1) all of tutored goal-directed actions can be regenerated robustly by setting each corresponding initial state obtained in the learning phase, that (2) the robot was able to create actional imaginaries of achieving the specified novel goal state with the consolidated memory at the critical regions of the initial state map, that (3) enactions of such generated plans brought potential instability in achieving the goal and however that (4) the reinforcements of some effective trials generated by chance gained the stability in regenerating those trials.

This result suggests that generation of novel or "creative" behaviors require two prerequisites for cognitive systems. One is that a good amount of sensory-motor experiences had been consolidated with self-organizing certain relational structures in distributed memories. The other is an existence of intrinsic cortical chaos which could enhance the criticality of the dynamic structure of the consolidated memory for generating diverse actional imaginaries. The continuous iterations of experiencing, consolidating and meditating would lead to developments of truly open-ended human-like intelligence.

Acknowledgment

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. A. R Luriia. *Higher cortical functions in man.* Basic Books, New York, 1966.
2. W. Penfield and J. Evans. The frontal lobe in man: A clinical study of maximum removals. *Brain*, (65):115–133, 1935.
3. K.M. Heilman and R.T Watson. *Intentional motor disorders.* Frontal Lobe Function and Dysfunction. Oxford University Press, New York, 1991.
4. S. Schoner and S. Kelso. Dynamic pattern generation in behavioral and neural systems. *Science*, 239:1513–1519, 1988.
5. R.D. Beer. A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72(1):173–215, 1995.
6. Jun Tani. Model-based learning for mobile robot navigation from the dynamical systems perspective. *IEEE Trans. on SMC (B)*, 26(3):421–436, 1996.
7. T. Ziemke. On 'parts' and 'wholes' of adaptive behavior: Functional modularity and diachronic structure in recurrent neural robot controllers. In *From animals to animats 6*, pages 171–180, Cambridge, MA, 2000. MIT Press.
8. C.A. Skarada and W.J. Freeman. Does the brain make chaos in order to make sense of the world? *Behavioral and Brain Sciences*, 10:161–165, 1987.
9. Jun Tani. An interpretation of the "self" from the dynamical systems perspective: a constructivist approach. *Journal of Consciousness Studies*, 5(5-6):516–42, 1998.
10. F.J. Varela, E. Thompson, and E. Rosch. *The embodied mind.* MIT Press, Cambridge, Mass., 1991.
11. J. Tani, R. Nishimoto, and R. Paine. Achieving "organic compositionality" through self-organization: Reviews on brain-inspired robotics experiments. *Neural Networks*, 21:584–603, 2008.
12. Marc Jeannerod. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17(2):187–245, 1994.
13. Ryunosuke Nishimoto, Jun Namikawa, and Jun Tani. Learning multiple goal-directed actions through self-organization of a dynamic neural network model: A humanoid robot experiment. *Adaptive Behavior*, 16(2-3):166–181, 2008.
14. C.L. Colby, J. Duhamel, and M.E. Goldberg. Ventral intraparietal area of the macaque: anatomic location and visual response properties. *Journal of Neurophysiology*, 69:902–914, 1993.
15. E. Eskandar and J. Assad. Dissociation of visual, motor and predictive signals in parietal cortex during visual guidance. *Nature Neuroscience*, 2:88–93, 1999.
16. H. Ehrsson, A. Fagergren, R. Johansson, and H. Forssberg. Evidence for the involvement of the posterior parietal cortex in coordination of fingertip forces for grasp stability in manipulation. *Journal of Neurophysiology*, 90:2978–2986, 2003.
17. Sugiko Imazu, Takeshi Sugio, Shigeki Tanaka, and Toshio Inui. Differences between actual and imagined usage of chopsticks: an fmri study. *Cortex*, 43(3):301–307, 2007.
18. Kenji Ogawa and Toshio Inui. Lateralization of the posterior parietal cortex for internal monitoring of self- versus externally generated movements. *Journal of Cognitive*

Neuroscience, 19:1827–1835, 2007.
19. Giacomo Rizzolatti, Luciano Fadiga, Vittorio Gallese, and Leonardo Fogassi. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2):131–141, 1996.
20. Ronald J. Williams and David Zipser. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1(2):270–280, 1989.
21. Kenji Doya and Shuji Yoshizawa. Memorizing oscillatory patterns in the analog neuron network. In *Proc. of 1989 int. joint conf. on neural networks*, volume 1, pages 27–32, Washington, D.C., 1989.
22. Yuichi Yamashita and Jun Tani. Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment. *PLoS Computational Biology*, Accepted.
23. G. Hesslow. Conscious thought as simulation of behaviour and perception. *Trends in Cog. Sci.*, 6(6):242–247, 2002.
24. Jukka Saarinen and Teuvo Kohonen. Self-organized formation of colour maps in a model cortex. *Perception*, 14(6):711–719, 1985.
25. Teuvo Kohonen. *Self-Organizing Maps*, volume 30 of *Springer Series in Information Sciences*. Springer, 2001.
26. David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. *Learning internal representations by error propagation*, volume 1 of *Parallel distributed processing: explorations in the microstructure of cognition*. MIT Press, Cambridge, 1986.
27. Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
28. WJ Freeman. *How Brains Make Up Their Minds*. Columbia University Press, 2001.
29. I. Tsuda, E. Koerner, and H. Shimizu. Memory dynamics in asynchronous neural networks. *Prog. Theor. Phys.*, 78:51–71, 1987.
30. K. Kaneko and I. Tsuda. Chaotic itinerancy. *Chaos*, 13(3):926–936, 2003.
31. S. Suzuki S. Sangawa Y. Ohmura K. Terada Y. Kuniyoshi, Y. Yorozu and A. Nagakubo. Emergence and development of embodied cognition: A constructivist approach using robots. *Progress in Brain Research*, 164:425–445, 2007.
32. P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Physical Review A*, 38:364–374, 1988.