Exploration Behavior in Shape Discrimination

Gentaro Morimoto and Takashi Ikegami

Graduate School of Arts and Sciences, The University of Tokyo
3-8-1 Komaba, Tokyo 153-8902, Japan
{genta,ikeg}@sacral.c.u-tokyo.ac.jp
Tel: +81-3-5454-4378; Fax: +81-3-5454-6541

Abstract. The motion of a model agent equipped with a plastic recurrent neural network to perform a shape discrimination task is simulated as an example of active perception. The task requires dynamic categorization of objects. The ability to discriminate shape is achieved by using an exploration behavior. We investigate the underlying mechanism of the deterministic exploratory behavior by showing the change of internal dynamics of the agent. The autonomy of the agent is also discussed.
Keywords: active perception, neural networks, plasticity, autonomous agent

1 Introduction

Gibson's theory of perception emphasizes the importance of movement of cognitive subjects (Gibson 1962). Perception is not merely information processing of sensory inputs, but an active process constructed from the subject's spontaneous movement and sensory inputs as a response to the movement from the environment. In other words, perception is impossible to discuss without the premise of embodied, situated subjects and their physical constraints.

Rethinking biological intelligence from the same viewpoint has been proposed by Walter (Walter 1950) and Braitenberg (Braitenberg 1984), and has recently been developed by many robotics researchers (Pfeifer and Scheier 1999, Nolfi and Floreano 2000). Simple vehicles equipped with relevant sensory-motor coordination(SMC) exhibit adaptive behavior without any symbolic representation in their brain. SMC should be a basis of artificial intelligence similar to the natural intelligence of living systems. However, when using this approach, it is difficult to achieve complex intelligent behavior, which can be achieved easily using classical computational artificial intelligence. There remains a considerable gap between models of embodied systems and truly intelligent cognitive systems.

We maintain that perception of artificial intelligent systems requires some form of dynamic representation, which is available independently of actual SMC. The representation is not symbolically predefined, but should be created while the subjects are interacting with objects. Dynamic categorization of patterns of interaction with objects is the essence of active perception and the first step toward symbolic computational behavior. The relationship between categorization and language ability has been discussed elsewhere (Lakoff 1987).

Currently, many researchers are endeavoring to create dynamic categorization models (Scheier and Pfeifer 1995, Nolfi and Morocco 2002, Beer 2003, Iizuka and Ikegami 2004). In order to extend vehicular models to demonstrate active perceptive behavior, we have introduced plasticity of weights in a recurrent neural network model (Morimoto and Ikegami 2004). It has been shown that our model agents generate dynamic categorization of objects due to this plasticity. The patterns of categorization vary greatly depending on the path of evolution. In this article, we investigate their exploration behavior and its stochastic nature from the viewpoint of internal dynamics. Finally, we discuss the autonomy of adaptive systems, which is found in their exploration behavior.

2 Model Description

2.1 Field and Task

The field is a two-dimensional discrete lattice space of 200×200 points. The boundary is periodic. Each point is in one of two states, empty(0) or occupied(1). An agent is placed on a point of the field and has a direction of the heading. It receives sensory inputs from the point it is on and the eight neighboring points. The positions of the sensory inputs can be distinguished relative to its direction. Therefore, nine bits of information in total can be used to decide the next movement. At every discrete time step the agent changes its position to a neighboring point and changes the direction of the heading according to the motor output. Motor output can be any of the three directions, namely, forward, left, or right.

There are objects that are classified either as rectangles or triangles in the field. This classification is determined by the global arrangement of the occupied points. Using upright and slanted edges, we designed two types of triangle and four types of rectangle. The environment contains 25 objects, of random type, direction and size. The size of each object is between 50 and 150 units in area.

The task of the agents is to pass more points in the rectangles and less in the triangles. This task implicitly requires categorization of the objects. Because they can see only 3 $\times$ 3 points in the field simultaneously, the categorization is required to be dynamic.

2.2 Network Architecture

An agent is equipped with a recurrent neural network to determine the next movement from the current state. The recurrent neural network is interpreted as a mapping of the internal variables, depending upon the inputs as parameters in terms of dynamical systems. Activations of context neurons represent the internal state that can be used to keep memory such as "I've turned left at the corner two time steps before."

Figure 1 depicts the internal network structure. The activations of neurons are updated internally as follows:

$$y_i(t) = g\big(\Sigma_j w_{ij} y'_j(t-1) + b_i\big), \tag{1}$$

$$g(x) = \big(1 + \exp\left(-\beta x\right)\big)^{-1}, \tag{2}$$

where $y'_i(t)$ is the actual activation of the ith neuron at time t, $y_i(t)$ is the value generated by the internal dynamics, w_{ij} is the weight of connection from the jth neuron to the ith neuron, and b_i is the bias of the ith neuron. The summation is taken over all neurons that have a connection with the ith neuron. $g(x)$ is the sigmoid function. As a result, the activations take a value from $(0, 1)$. β is the nonlinearity coefficient and is set to 1.0 in this paper.

Actual activations of input neurons are modified as follows:

$$y'_i(t) = (1-\mu)s_i(t) + \mu y_i(t), \tag{3}$$

where s_i is the raw sensory input and takes one of two values, (0)(empty) or (1)(occupied). For context and output neurons $y'(t) = y(t)$. The meaning of (3) is that the actual activations of input neurons are modified by the value that is generated from the internal dynamics. In this paper, we used $\mu = 0.3$. Therefore, the activation of an input neuron takes a value from $(0, 0.3)$ or $(0.7, 1.0)$, depending on the state of the corresponding position in the field. When the value from internal dynamics is close to the sensory inputs, actual activations of input neurons become clearly distinguishable.

In addition, we introduced plasticity of weights. Weights from input neurons and context neurons into input neurons change during the interaction with the environment so that input neurons can play a role similar to prediction. These weights are updated as:

$$\Delta w_{ij}(t) = \eta\big(s_i(t) - y_i(t)\big)y'_j(t-1), \tag{4}$$

where $\eta (= 0.01)$ is the learning rate. By introducing plasticity in this way, the value y_i corresponding to the input neurons, which is generated from the internal dynamics, generally has the tendency to approach s_i.

3 Results

We produced an agent that has relevant SMC to perform the shape discrimination task using a genetic algorithm. The initial values of weights were encoded and the agent's performance of the task was evaluated. Figure 2 shows a trajectory of the agent in an environment. The agent wanders between triangles and after finding a square object, it changes its motion to examine the shape of the object and finally starts to fill in the object. The right part of Figure 2 shows the time series of activations of output neurons. The pattern of movement is classified into three phases. The first period (0–650) is the wandering phase, the second period (650–1000) is the examining phase, and the third period (1000–1500) is the filling phase. In each phase, the actual value of the activations slowly changes because of the plasticity. For example, in the wandering phase, a period doubling bifurcation occurs at approximately 300 time steps.

Figure 3 shows the difference between three phases in detail. The dynamic patterns of output neurons indirectly show the patterns of the agent's actual movement. In the left part of the figure, the internal states

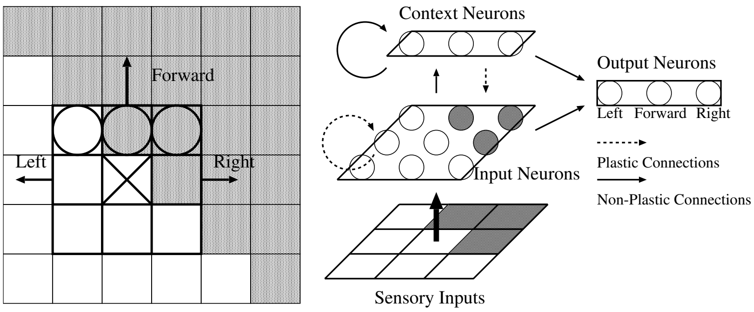

Figure 1. An agent with neural network architecture is situated in the environment. The architecture consists of nine sensory inputs that are connected to corresponding input neurons. Three context neurons are used to keep short-term memory in unit cubic space. Movements are decided by the most activated neuron of the three output neurons related to moving forward, left or right. Connections to input neurons are plastic and updated according to the sensory inputs.

of the agent are plotted in three-dimensional context neuronal space. The difference between phases is also represented in the area of internal states.

Plasticity introduced in our model enables deterministic unforeseen exploratory behavior, which make it possible to achieve higher performance than is the case without plasticity. The network weights and temporal internal states create a dynamic memory of interaction. This enables the switching of movement pattern between three phases.

However, the switching is not discrete and clear. It is not as if there are threshold neurons to represent the current phase. Action plans are generated in parallel, and one of them is selected by the interaction with the environment. Figure 4 shows another example of a trajectory and time series of activations of output neurons by the same agent in a different environment. It can be seen that switching between phases is not always one directional. The learning pathway in the weight space is dependent on how interaction with the environment has occurred.

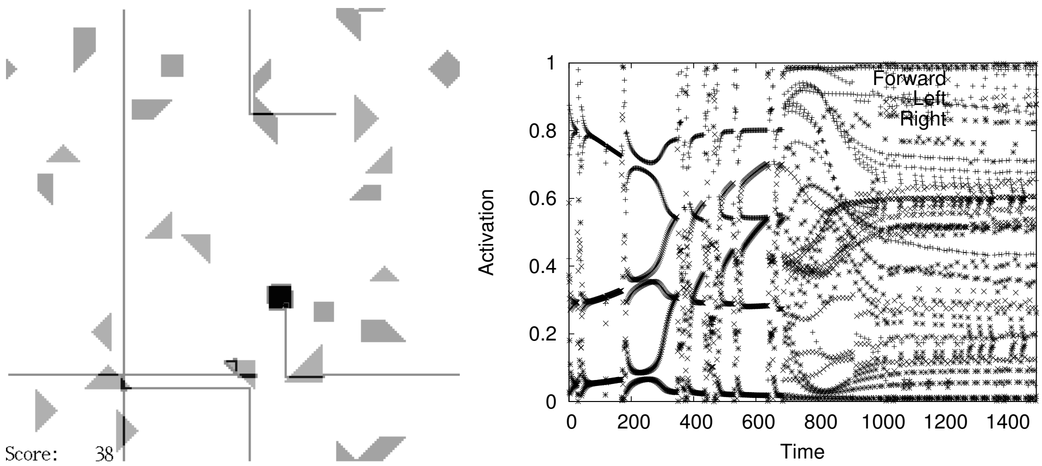

Figure 2. An example of movement of an agent. The left is the environment and the trajectory. First, the agent moves straight in the empty zone and changes direction when it encounters an object in the field. Second, when it encounters a candidate, a rectangular object, it periodically moves back and forth between the edges to examine the shape of the object. Finally, it starts changing its direction nonperiodically to fill the object and eventually passes all the points inside the square. The right is the time series of activations of the output neurons. Period 0–650 is the wandering phase, period 650–1000 is the examining phase, and period 1000–1500 is the filling phase.

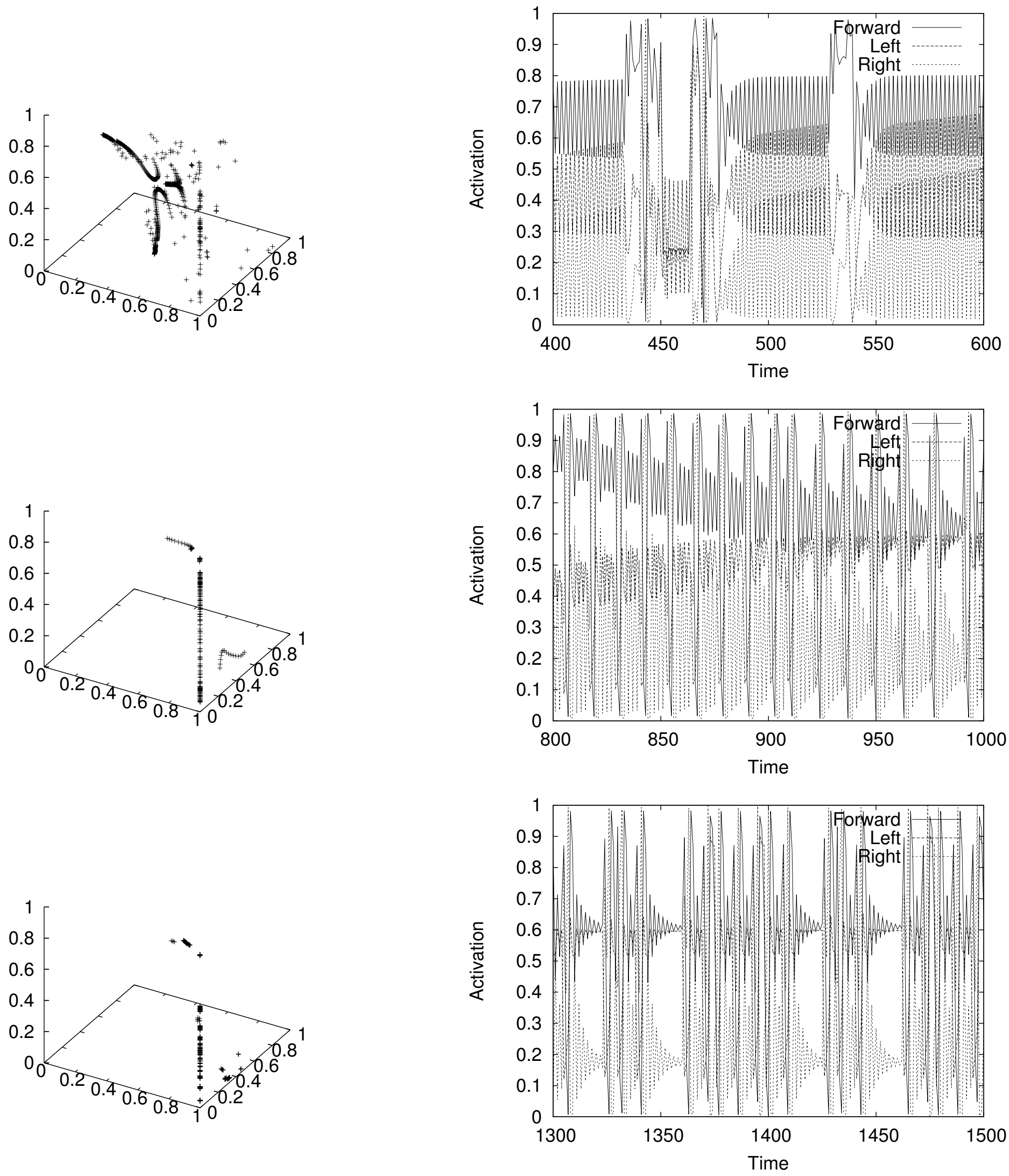

Figure 3. The difference of dynamics in the three phases of movement (top: wandering; middle: examining; bottom: filling). The left figures show the values of the internal states at each time step in the context neuronal space. The right figures show the time series of activations of the output neurons. In the wandering phase, the internal dynamics show the oscillation with period two and moving forward is always selected in an empty zone. When an object is visible, a change of state, or turning left or right is sometimes selected. In the examining phase, the agent moves forward inside the object and changes direction at the edge by turning twice. It periodically moves between the facing edges, back and forth. After a while, movement changes to the filling phase. In the filling phase, the agent changes direction nonperiodically at the edge or even inside the object. The unforeseen turning due to plasticity enables filling of all points of the square.

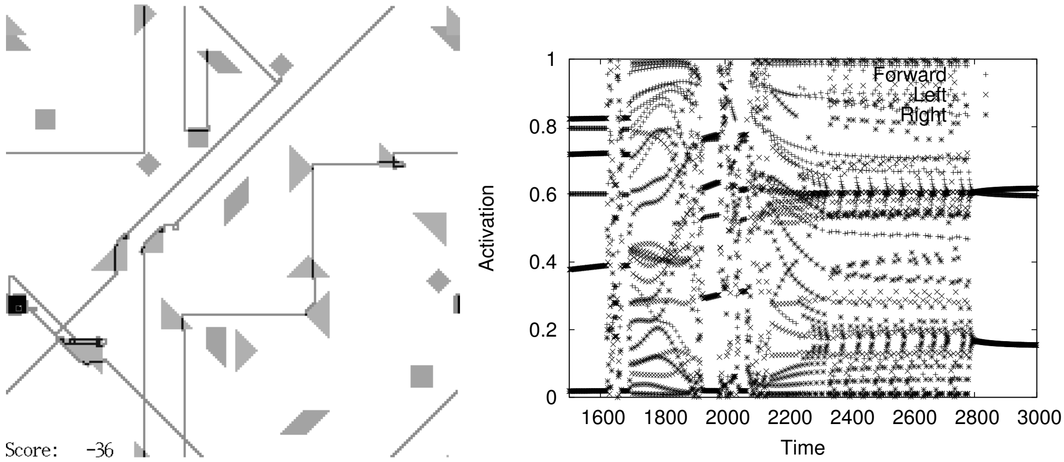

Figure 4. An example of movement of the same agent in another environment. The wandering phase takes longer than for the former example. In the latter period of the wandering phase the agent moves obliquely. After encountering a square object, its movement changes to the examining phase, but accidentally leaves the object. It happens to reenter the object and succeeds in switching to the filling phase. The memory of touching a candidate square is retained.

4 Discussion

4.1 Effect of Plasticity on Exploration Behavior

We introduced plasticity of the network weights assuming a prediction-like role of the input layer as a criterion of Hebbian learning. The basis of the idea of plasticity is that the same sensory input stimuli should be able to create different neural dynamics and activate different movements depending upon earlier experience. The prediction-like role in the input layer tries to match the intrinsic dynamics to extrinsic dynamics. One of the benefits of introducing prediction-like plasticity is that we can suppose stand-alone internal dynamics without any sensory inputs, substituting 1.0 for μ. We can find chaos and bifurcation in the autonomous internal dynamics. The instability of the internal dynamics is a source of the exploration behavior.

Without plasticity the movement of the model agent became deterministically dependent on temporal input patterns and the saturated value of the performance became lower in the genetic algorithm. Theoretically, it is possible to realize the historical effect of the movement by a simple recurrent neural network without plasticity. It may require greater computation resources (e.g., more context and hidden neurons) and be more difficult to evolve.

Prediction may not be crucial for adaptive behavior. We suppose the essential role of plasticity is to introduce slower time-scale dynamics compared with the time-scale of navigation. The slow dynamics enhance the variety of behavior in the lifetime of the agents and makes it possible for them to adapt to various environments. The adaptability due to mixing different time-scale dynamics inevitably exhibits exploration behavior.

4.2 Action Selection and Autonomy

Although what is examined here is only one example of active perception, this mechanism naturally constructs a loose structural coupling between the agent and the environment. The interaction is highly dependent on history and the realized trajectory is one of many possible trajectories. In some points of the trajectory the action selected was from almost equally possible actions. Figure 5 shows the possible trajectories. It is shown that the branching of trajectories does not occur randomly, but has a structure that depends on the movement phases. The whole branching tree of trajectories comprises a bundle of possible actions and is illustrative of the manner of interaction between the agent and the environment.

The autonomy of adaptive systems is characterized by the bundle of unrealized possible actions and the lack of reason for action selection. This viewpoint of possible world structure has been proposed in the context of strategy selection by game players (Ikegami and Taiji 1998). Personal experience generates a tree of what was possible in the past and can affect action selection in the future. What is important in the properties of autonomous agents is to leave room for possible actions from the past. This can be a compromise between a deterministic worldview and subjective autonomy.

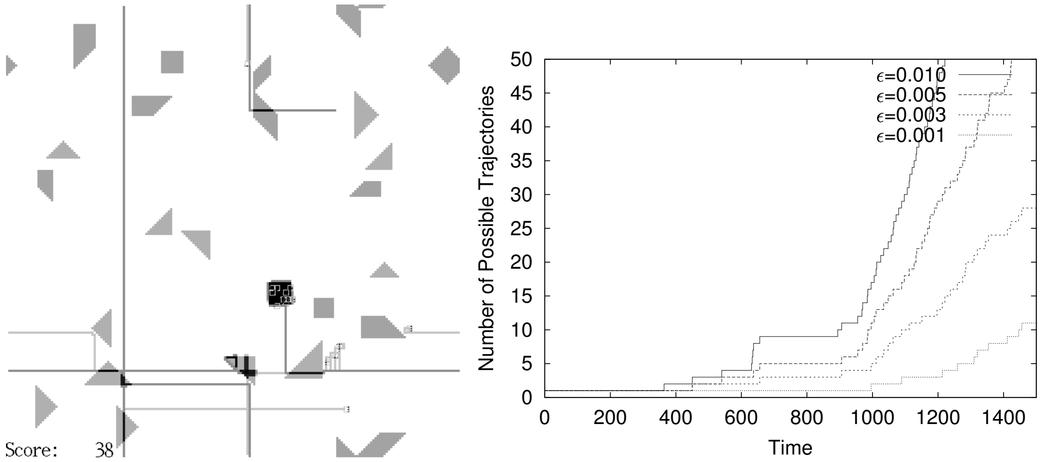

Figure 5. If the difference between the most and the second-most activated neurons is small, there is the possibility of another trajectory in the presence of some noise or a small difference in the history of interaction. If the difference is smaller than ϵ, the simulation of movement of the agents is duplicated and the possible trajectory is traced for 100 time steps in order to consider the bundle of possible actions. The left figure shows all possible trajectories when $\epsilon = 0.01$. The right figure shows the number of possible trajectories branched from the real trajectory as functions of ϵ. In the wandering phase, branching often occurs at the edge of an object. In the examining phase, branching seldom occurs. In the filling phase, branching occurs randomly and the number of possible trajectories grows linearly with time. All of the possible trajectories that branched in the filling phase remain inside the square.

Acknowledgments

This work is partially supported by Grant-in aid (No. 09640454) from the Ministry of Education, Science, Sports and Culture and from The 21st Century COE (Center of Excellence) program (Research Center for Integrated Science) of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

Beer, R. D. (2003), “The dynamics of active categorical perception in an evolved model agent,” *Adaptive Behavior*, vol. 11(4), pp. 209–243.

Braitenberg, V. (1984), *Vehicles*, MIT Press, Cambridge MA.

Gibson, J. J. (1962), “Observations on active touch,” *Psychological Review*, vol. 69, pp. 477–491.

Iizuka, H. and Ikegami, T. (2004), “Simulating autonomous coupling in discrimination of light frequencies,” *Connection Science*, vol. 16(4), in press.

Ikegami, T. and Taiji, M. (1998), “Structures of possible worlds in a game of players with internal models,” *Acta Polytechnica Scandinavica Ma.*, vol. 91, pp. 283–292.

Lakoff, G. (1987), *Women, Fire, and Dangerous Things*, The University of Chicago Press, Chicago, IL.

Morimoto, G. and Ikegami, T. (2004), “Evolution of plastic sensory-motor coupling and dynamic categorization,” *Proceedings of Artificial Life IX*, pp. 188–193.

Nolfi, S. and Floreano, D. (2000), *Evolutionary Robotics*, MIT Press, Cambridge, MA.

Nolfi, S. and Morocco, D. (2002), “Active perception: A sensorimotor account of object categorization,” *From Animals to Animats VII; Proceedings of the International Conference on Simulation of Adaptive Behavior*, pp. 266–271.

Pfeifer R. and Scheier C. (1999), *Understanding Intelligence*, MIT Press, Cambridge, MA.

Scheier, C. and Pfeifer R. (1995), “Classification as sensory-motor coordination: A case study on autonomous agents,” *Proceedings of the 3rd European Conference on Artificial Life*, pp. 657–667.

Walter, W. G. (1950), “An imitation of life,” *Scientific American*, vol. 182(5), pp. 42–45.