International Symposium on Nonlinear Theory and its Applications
Xi'an, PRC, October 7–11, 2002

Learning to the Movement Direction Selection Function without Teacher's Signal

Hiroki TAMURA, Zheng TANG, Weidong SUN and Masahiro ISHII

Department of Intellectual Information Systems Engineering
Toyama University
3190 Gofuku, Toyama-shi, Japan
Phone:+81-76-445-6886, Fax:+81-76-445-6752
Email: tamura@iis.toyama-u.ac.jp

1. Introduction

The neuron has the characteristic of recating to a specific stimulus[1]. However, it is not clear how the neuron learn the function. The neuron model which authors proposed is a neuron model with consideration to the interaction of a dendrite-like projection and difference of time of inputs[2,3]. We have succeeded in making the movement direction selection function (one-dimension and two-dimension) learn by this neuron model[2]. Although the teacher's signal makes learning possible, it does not know whether such a teacher signal exists in actual biology.

In this paper, a network model(column) which learns the movement direction selection function without a teacher's signal is proposed. The network model which consists of the plurality of the neuron model we proposed previously is described. Teacher's signal is automatically created with the simple rule in the network model. The simulation of the four direction selection(→, ←, ↓ , ↑) on two-dimension is performed using the network model and, the results show the validity and learning capability of this network model.

2. Network model

The network model proposed in this paper performs three processings. The 1st processing is input-processed. The 2nd processing performs a neuron simple substance. The 3rd processing deals with two or more neuron model in a column.

2.1. The 1st processing(input)

The input is shown to the network model when a stimulus is given to the specific domain on the retina. The neuron model is made to realize the movement direction selection function by giving difference of time of inputs. The input which received the stimulus on the retina is taking into consideration the input of two types, the fast input which travels to a neuron model through fast pathway, and the slow input which travels to a neuron model through slow pathway. In this paper, inputs are made into the domain of 3x3, and the input which the neuron model receives is set to 18 from the fast inputs and the slow inputs.

2.2. The 2nd processing(neuron model)

Fig.1 is the single neuron model which we proposed. This neuron model takes into consideration the interaction between the synapses on a dendrite-like projection, and the difference of time of inputs. Fig.1 shows that Connection Layer is synapse , AND layer is the interaction between the synapses inputted into a dendrite projection ,OR Layer is the branch of a dendrite projection and Soma Layer is Soma of neuron. At this neuron model, the interaction between the synapses on a dendrite-like projection is realized by AND Layer and OR Layer. Moreover, the input of two types, the fast inputs and the slow inputs, is inputted into this neuron model by dendrite-like projection of one from the 1st processing through synapse combination to an input, respectively. The difference of time of inputs can be taken into consideration from this capability. Soma layer receives the input from a denderite-like projection. Soma layer has threshold and an internal state and the input from a dendrite-like projection is accumulated at an internal state. Soma layer spreads and when the threshold is exceeded, a neuron model ignites. The neuron model which ignited at once presupposed that it does not ignite however high an internal state may become during a certain fixed period(refractory period).

Input Layer

An input layer sets the input from the exterior to $x_i(i = 0, 1, \cdots, n)$. The input from fast pathway

$$X_i(t) = x_i(t) \quad (1)$$

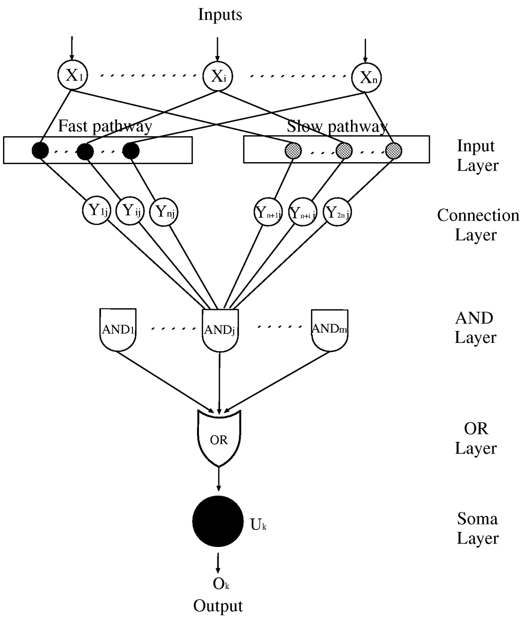

Figure 1: The neuron model.

and, the input from slow pathway

$$X_{n+i}(t) \quad = \quad x_i(t - \Delta t) \tag{2}$$

It carries out and two inputs are told to a connection layer. Here, t expresses the time change of an input. Δt is fixed to 1 in this paper. It is an input from a late course by this. It will be in the input state in front of one.

Connection Layer

Each node in this layer is a connection function and yields the direct connection(excitatory synapse) or the inversed connection(inhibitory synapse) or the 1-constant connection or the 0-constant connection, to corresponding AND gates.

The connection function can be described by a one-input one-output

$$Y_{ij} \quad = \quad f(u_{ij}) = \frac{1}{1 + e^{-hu_{ij}}} \tag{3}$$

where $u_{ij} = w_{ij}X_i - \theta_{ij}(1 - X_i)$, W_{ij}, θ_{ij} are connection parameters and h is a positive constant. Since the input $X_i(0 \leq i \leq 2n)$ has only two constants, 0 and 1, for a large h , there are only four cases for different weights W_{ij} and threshold θ_{ij} as follows.

1. 1-constant connection
   $0 < w_{ij}, 0 > \theta_{ij}$, corresponds to the 1-constant connection.
2. 0-constant connection
   $w_{ij} < 0, 0 < \theta_{ij}$, produces the same 0-constant connection.
3. inhibitory synapse
   $w_{ij} < 0, 0 > \theta_{ij} < 0$, gives the inversed connection.
4. excitatory synapse
   $0 < w_{ij}, 0 < \theta_{ij}$ presents the direct connection.

Therefore, the connection states can be described by four states of connection.

AND layer

A node in layer 2 corresponds to the AND operation. The AND operation of the input variable yields the value 1 if and only if all input variables are simultaneously 1. In order to do learning with gradient descent, derivatives of the AND function are needed. However, the derivatives of the AND function do not exist. Here, we use a softminmum operator.

$$f_{min}(x_1, x_2) = \frac{x_1 e^{-ux_1} + x_2 e^{-ux_2}}{e^{-ux_1} + e^{-ux_2}} \tag{4}$$

$$AND_j = f_{min}(Y_{11} \cdots Y_{ij} \cdots Y_{2nj}) \tag{5}$$

where μ is a positive constant. The softmin function produces the same result as the AND operation in the limit.

$$\lim_{u \to \infty} f_{min}(x_1, x_2) = \min\{x_1, x_2\} \tag{6}$$

Furthermore, the function is also clearly differential.

OR layer

A node in layer 3 is OR function. OR function is expressed with a soft-maximum function like AND function.

$$f_{max}(x_1, x_2) = \frac{x_1 e^{vx_1} + x_2 e^{vx_2}}{e^{vx_1} + e^{vx_2}} \tag{7}$$

$$OR = f_{max}(AND_1 \cdots AND_j \cdots AND_m) \tag{8}$$

where v is a positive constant. The softmax function produces a same result as the OR operation in the limit.

$$\lim_{v \to \infty} f_{max}(x_1, x_2) = \max\{x_1, x_2\} \tag{9}$$

A soft-minimum function and a soft-maximum function can be learned by differentiation being possible and correcting connection parameter w_{ij} of a connection function, and θ_{ij} according to the back propagation-like algorithm by the continuous function.

Soma Layer

Soma layer receives the input from OR Layer. Soma layer has threshold(θ_{soma_k}) and an internal state(U_k).

The input from OR layer is added to an internal state. When the internal state about Soma layer exceeded threshold, a neuron model ignites($O_k = 1$). The neuron model which ignited at once presupposed that it does not ignite howmuch high an internal state may become during a certain fixed period (Timer > 0). It can avoid becoming a neuron model which ignites to any patterns by that. An internal state is expressed with Eq.(10).

$$U_k = \sum_{t=1} OR(t) - \theta_{soma_k} \tag{10}$$

and the conditions of ignition are expressed with Eq.(11).

$$O_k = \begin{cases} 1 & (U_k \geq 0 \text{ and Timer} = 0) \\ 0 & (U_k \geq 0 \text{ and Timer} > 0) \\ 0 & (U_k < 0) \end{cases} \tag{11}$$

2.3. The 3rd processing(column)

The network model proposed in this paper is shown in Fig.2. The neuron model of Fig.2 has the structure which was shown in Fig.1, respectively. If any one neuron model in a column ignites, I(t) ($\sum_{k=1}^{N} W_T O_k(t)$) will be received. In this paper, W_T is set to 1. Connection parameter w_{ij} and θ_{ij} are corrected so that it may not ignite noting that the neuron which received I(t) does not need to ignite. Moreover, the neuron model which ignited corrects connection parameter w_{ij} and θ_{ij} so that it may ignite more to the ignited input's pattern. That is, it becomes the role of a teacher's signal when the information on ignition of the neuron model in a column learns. Moreover, when not all neuron model have ignited in a column, connection parameter w_{ij} and θ_{ij} are corrected so that all the neuron model in a column may make an output increase. That is, the learn method of a neuron model is decided in the 3rd processing. These processings about neuron model k are realized by Eq.(12,13).

$$\Delta w_{ij}(t) = -T \frac{\partial OR(t)}{\partial w_{ij}(t)} \tag{12}$$

$$\Delta \theta_{ij}(t) = -T \frac{\partial E(t)}{\partial \theta_{ij}(t)} \tag{13}$$

T is decided to Eq.(14) as conditions of learn.

$$T = \begin{cases} \eta 1(OR(t) - 1) & : (O_k(t) = 1\ ) \\ \eta 2(OR(t) - 1) & : (O_k(t) = 0 \text{ and } I(t) = 0) \\ \eta 3(OR(t) - 0) & : (O_k(t) = 0 \text{ and } I(t) > 0) \end{cases} \tag{14}$$

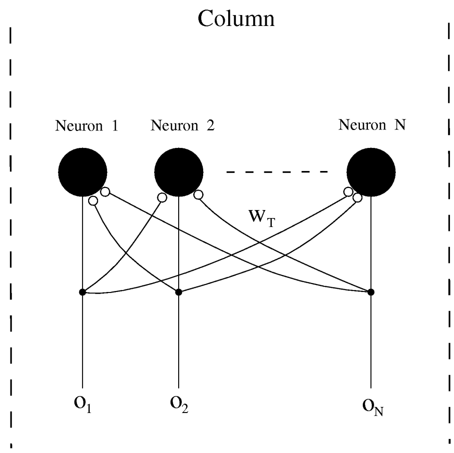

Figure 2: The network model(Column).

Table 1: Learning Parameter.

The learning constant $\eta 1$	0.3
The learning constant $\eta 2$	0.2
The learning constant $\eta 3$	0.5
The learning constant θ_{soma_k}	1.0
The positive constant h of sigmoid function	5.0
The positive constant u of softmin function	5.0
The positive constant v of softmax function	5.0

In this network model, learn becomes having no teacher's signal and possible about the movement direction selection function by performing the 3rd processing from the 1st processing.

3. Simulation

We simulated that the movement direction selection function of four directions($\rightarrow, \leftarrow, \downarrow, \uparrow$) shown in Fig.3 learn by the network model to propose. The value of the connection parameter(w_{ij},θ_{ij}) which determines the joint state of synapse connection as an initial state of each neuron model was chosen at random in the range of $-1.5 < w_{ij}, \theta_{ij} < 1.5$. The parameters of Table 1 are used in the simulation. The network model was made to give and learn at random the four movement direction patterns. The number of the neuron models which form a column is set to 10 and the number of dendrite-like projections which a neuron model has is set to 10. In this simulation, once neuron model ignites, neuron model does not ignite until input's pattern will be inputted 3 times(Timer = 3). The movement direction pattern is made to input 10000 times.

From simulation results, the internal state (U_k) of the neuron model to the inputs pattern of the output of the neuron before learning and the neuron after learning is

Table 2: The internal state(U) and Output(O) of neuron model before learning(above),after learning(below).

	Neuron number	1	2	3	4	5	6	7	8	9	10
pattern1	output	0	0	0	0	0	0	0	0	0	0
	(U)	(0.295)	(0.297)	(0.228)	(0.290)	(0.319)	(0.274)	(0.279)	(0.311)	(0.308)	(0.305)
pattern2	output	0	0	0	0	0	0	0	0	0	0
	(U)	(0.296)	(0.280)	(0.222)	(0.273)	(0.312)	(0.271)	(0.249)	(0.288)	(0.295)	(0.304)
pattern3	output	0	0	0	0	0	0	0	0	0	0
	(U)	(0.297)	(0.286)	(0.231)	(0.292)	(0.329)	(0.281)	(0.254)	(0.311)	(0.318)	(0.324)
pattern4	output	0	0	0	0	0	0	0	0	0	0
	(U)	(0.303)	(0.283)	(0.237)	(0.296)	(0.334)	(0.278)	(0.258)	(0.312)	(0.311)	(0.328)

	Neuron number	1	2	3	4	5	6	7	8	9	10
pattern1	output	0	0	0	0	0	0	0	**1**	0	**1**
	(U)	(0.538)	(0.590)	(0.498)	(0.532)	(0.545)	(0.557)	(0.773)	(2.332)	(0.568)	(2.887)
pattern2	output	0	0	0	0	**1**	0	0	0	0	0
	(U)	(0.546)	(0.582)	(0.519)	(0.512)	(2.919)	(0.583)	(0.707)	(0.710)	(0.599)	(0.557)
pattern3	output	**1**	0	**1**	**1**	0	0	0	0	0	0
	(U)	(2.888)	(0.559)	(2.886)	(2.884)	(0.559)	(0.590)	(0.870)	(0.561)	(0.604)	(0.589)
pattern4	output	0	0	0	0	0	0	**1**	0	0	0
	(U)	(0.549)	(0.592)	(0.598)	(0.567)	(0.559)	(0.563)	(2.296)	(0.600)	(0.571)	(0.562)

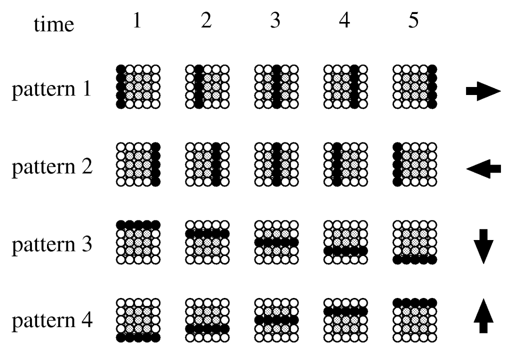

Figure 3: Input patterns.

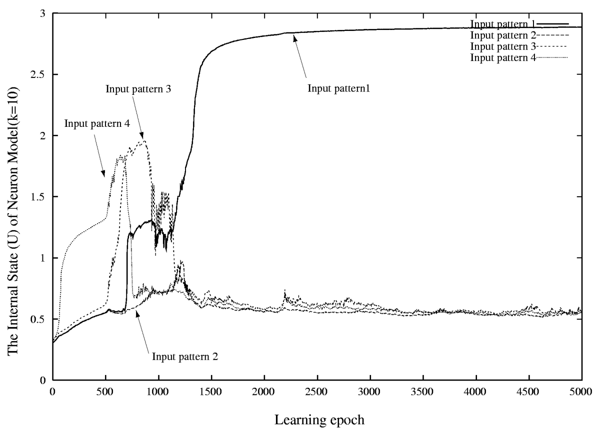

Figure 4: Learning to Inputs Pattern 1 (neuron 10).

shown in Table 2. The neuron model in a column different to each inputs pattern came to ignite after learning to the neuron model having reacted to no input patterns before learning, as shown in Table 2. For example, The change of the internal state to each inputs pattern of the neuron model(10:learn to input's pattern 1 (→)) is shown in Fig.4. The neuron model(10) reacts only to inputs pattern 1.

It suggests that the network model may be possible to learn the solution to the movement direction selection function without teacher's signal.

4. Conclusions

We proposed the network model which can learn the movement direction selection function without a teacher's signal. By determining the law of learning according to the state of two or more neuron models in a column, even if this network model has no teacher's signal, it can be learned. Simulation results showed that the network model proposed learned the movement direction selection function.

References

[1] D.O.Hebb,: "The organization of behaviour", Wiley,New York,1949.

[2] H. Tamura, Z. Tang, O. Ishizuka, K. Tanno,: "Directionally selective cells have a δ -like morphology", NOLTA, Hawaii, U.S.A, pp215-218,1999.

[3] Z. Tang, H. Tamura, M. Kuratsu, O. Ishizuka, K. Tanno,: "A model of the neuron based on dendrite mechanisms", Electronics and Communications in Japan, Vol.84, No.8:pp11-24,2001.