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

Characteristics of Inverse Delayed Model for Neural Computation

Koji NAKAJIMA† and Yoshihiro HAYAKAWA‡

Laboratory for Electronic Intelligent Systems, Research Institute of Electrical Communication, Tohoku University
Katahira 2-1-1, Aobaku, Sendai, 980-8577, Japan
Phone:+81-22-217-5558, Fax:+81-22-217-5558
Email: †hello@nakajima.riec.tohoku.ac.jp, ‡asaka@riec.tohoku.ac.jp

1. Introduction

Many models of neural cells and networks have been proposed and investigated. The typical cell models are, for examples, Hodgkin-Huxley model [1] based on physiological assumptions, FitzHugh's BVP model [2] which is a projection of the Hodgkin-Huxley model on a phase plane, McCulloch-Pitts [3] and Caianiello [4] models for analyses of network functions, the continuous-valued units model [5], [6], Aihara model [7] for chaotic behaviors, etc. And there are many studies for layer and full connection structures, learning functions, statistical methods, etc. [8]

The most of existing neural network models interpret the biological behaviours in terms of time averaging techniques, rather than directly emulating them. It is a resonable simplification to handle macroscopically the collective computational properties of the networks. On the other hand, the direct emulations are main concerns to the single or microscopic neuron models. It seems there are some gaps between these macroscopic and microscopic models. Particularly, the ability of excitation which is an important characteristic for neurons is not fully included in the macroscopic models. A normal neuron, which, however, seldom shows spontaneous excitations, is ordinarily in a resting state. And it fires depending on inputs from other neurons and external stimulations. It is a fundamental condition for the neurons which are in the critical state to be able to fire. The condition is also considered to have many rolls for the functions of neural networks. It is an elementary assumption to have two states resting and excitation for neurons in the past network models where the output of neuron units corresponds to a short-term average firing rate. And hence, it seems few discussions have been presented what rolls the neuron in the critical state or the ability of excitation has on the network actions.

In this paper, we present the novel model (Inverse Delayed Model) to represent networks with neurons in the critical state, and discuss the fundamental characteristics of the ID model including stimulus responses and an energy function of the the network. The ID model has a simple structure, however, it can clear up the relation between the macroscopic and the microscopic models, because it includes both of the BVP and the Hopfield models.

2. Basic Equations

Real neurons have continuous input-output relations and integrative time delays, and hence, the time evolution of the state of such systems should be represented by differential equations. The conventional continuous-valued unit model is represented by

$$\tau \frac{du}{dt} = \sum_j W_j x_j - u, \tag{1}$$

$$x = f(u), \tag{2}$$

where u, x, W_j, and $f(u)$ are an internal state of a certain unit, an output, a connection strength from j-th unit, and an output function, respectively, which is a sigmoid function in conventional models. The conversion from u to x has negligible response time in comparison with τ in Eq.(1). However, the conversion time should be taken into consideration in general cases [5], [13] under the condition that it is much smaller than the τ. Let us consider the following equations instead of Eq.(2),

$$\tau_x \frac{dx}{dt} = f(u) - x, \tag{3}$$

or

$$\tau_x \frac{dx}{dt} = u - g(x), \tag{4}$$

where $\tau_x << \tau$ and $g = f^{-1}$ is defined as the inverse input-output relation [6]. Equations (3) or (4) is equal to Eq. (2) when $dx/dt = 0$, and the three equations are equivalent at the limit $\tau_x \to 0$. It is convenient to use Eq. (4) rather than Eq. (3) when a hysteretic (S type) function is considered for $f(u)$ [9], [10], [11] instead of a monotonic increasing sigmoid-function. Because the inverse function g is still a single value (N type) function even when f is hysteretic and hence a multivalue function. From these points of view, our Inverse Delayed (ID) Model is represented by Eqs. (1) and (4). The time delay of the input-output relation and the N type function for g lead to negative resistance characteristics [12] which give an ability of firing in each unit. The ID model enable us to make continuous discussion from the case of a monotonic increasing sigmoid-function to a hysteretic function for f. More, the ID model, which is obtained from the extension of the macroscopic model, can be converted into the BVP model which is one of the physiological microscopic model, if we assume a specific g function and certain parameters. And it is also easy to discuss the network behaviour according to the concept of the energy functions as shown in the following section. Let the self-connection strength and inputs from other units be W and θ, respectively, which also represents external stimuli and bias inputs. We assume W and θ are constant only in the discussion for the behaviour of a single unit. Eq. (1) is rewritten

$$\tau\frac{du}{dt} = Wx + \theta - u. \tag{5}$$

If we put $\tau = c/b, \tau_x = 1/c, W = -1/b, \theta = a/b, g(x) = x^3/3 - x$ in Eqs. (4) and (5), the ID model is reduced to the BVP model without stimulus intensity z [2]. Eq. (4) in the BVP model contains stimulus intensity z, on the other hand, θ in Eq. (5) of ordinary macroscopic models represents also the stimulus intensity.

The features of the BVP model in view of the ID model are the followings, 1) It has negative self-connection whose absolute value is larger than 1, because $0 < b < 1$ as stated in Ref. [2], 2) Its inverse output function $g(x)$ is $x^3/3 - x$ where $x^3/3$ and $-x$ give a loose confinement near $x = 0$ to outputs and a negative resistance, respectively, 3) $\tau_x/\tau = b/c^2 < 1$ in the BVP model as stated in Ref. [2].

3. Negative Resistance

In order to obtain the expression of the negative resistance, Eq. (4) is differentiated with time, and substitutions of Eqs. (5) and (4) lead to the following differential equation,

$$\tau_x\frac{d^2x}{dt^2} + \eta\frac{dx}{dt} = -\frac{\partial U}{\partial x} \tag{6}$$

$$\eta = \frac{dg(x)}{dx} + \frac{\tau_x}{\tau} \tag{7}$$

$$\frac{\partial U}{\partial x} = \frac{1}{\tau}\left\{g(x) - Wx - \theta\right\} \tag{8}$$

Equation (6) represents a particle motion on the potential U which is obtained by the integration of Eq. (8). The first and the second terms on the left hand side of Eq. (6) corresponds to a inertia and a friction terms, respectively. η is a coefficient of the friction and a nonlinear function of the output x.
When $g(x)$ is a monotonic increasing function, for an example x^3 or $arctanh(x)$, U has a single well structure for $W \leq 0$ and it has a double well (two minima) structure for $W > 0$. If $g(x)$ is an odd function and $\theta = 0$, $x = 0$ is a only minimum point in the single well structure, and $x = 0$ is an unstable equilibrium point in the double well structure.

The coefficient η of the friction is positive when $g(x)$ is a monotonic increasing function. However, it becomes negative if $g(x)$ has a N shape which gives a S shape to the output function f. The hysteretic (S shape) output function f does not always mean a hysteresis unit which is due to a double well potential. The cause of the double well structure does not only depend on the shape of output function but also the self-connection strength W, which does not cause negative resistance as Eq. (7) shows. Therefore, the unit as the BVP model does not always have hysteretic characteristics depending on the value of the self-connection strength, even if it has g with a N shape and a negative resistance.

Let us discuss two examples, $g(x) = x^3/3 - x$ and $g(x) = (1/\beta)arctanh(x)$ ($f(u) = tanh(\beta u)$) which is often used for the discussion of network collective computations. In the first example which is the BVP model, the coefficient η_{BVP} and the negative resistance region on x are

$$\eta_{BVP} = x^2 - \left(1 - \frac{\tau_x}{\tau}\right), \tag{9}$$

$$|x| < \sqrt{1 - \frac{\tau_x}{\tau}}. \tag{10}$$

And the potential U_{BVP} is

$$U_{BVP} = \frac{1}{\tau}\left\{\frac{x^4}{12} - \frac{(1+W)}{2}x^2 - \theta x\right\} \tag{11}$$

The self-connection strength W in the BVP model is smaller than -1, and hence, the unit does not have hysteretic characteristics because of the only one equilibrium point on the potential. The unit turns to a continuous firing state when $\theta = 0$, because the equilibrium point $x = 0$ is in the negative resistance region, and the unit is unstable according to the conventional small signal analysis around the point. The unit is in the resting state when $|\theta| > \{\frac{1}{3}(1 - \frac{\tau_x}{\tau}) - (1 + W)\}\sqrt{1 - \frac{\tau_x}{\tau}}$ where $W < -1$ in the BVP model, because $\eta_{BVP} = 0$ at $x = \sqrt{1 - \frac{\tau_x}{\tau}}$. When $W > -1$ and $\theta = 0$, the equilibrium points are $x = 0$ (unstable) and $x = \pm\sqrt{3(1 + W)}$ which are two stable equilibrium points if $W > -\frac{1}{3}(2 + \frac{\tau_x}{\tau})$, and hence, the unit is in the resting state at the points.

In the second example where $g(x) = (1/\beta)arctanh(x), (\beta > 0)$, if we change the inverse output function for $g(x) = g_0(x) - g_1(x) = (1/\beta)arctanh(x) - x$ similar to the BVP model, the coefficient η_a and the negative resistance region on x are

$$\eta_a = \frac{1}{\beta(1-x^2)} - \left(1 - \frac{\tau_x}{\tau}\right), \tag{12}$$

$$|x| < \sqrt{1 - \frac{\tau}{\beta(\tau - \tau_x)}}, \tag{13}$$

where $\beta > \tau/(\tau - \tau_x) > 1$, because $\tau >> \tau_x$. Therefore, this model has no negative resistance if $\tau < \tau_x$ which, however, is against the assumption in the previous section. And the potential U_a is

$$U_a = \frac{1}{\tau}\left\{\frac{x}{\beta}arctanh(x) + \frac{1}{2\beta}log\left(1 - x^2\right) - \frac{1+W}{2}x^2 - \theta x\right\} \tag{14}$$

The potential has a deep well shape, and hence, it is expected that the characteristics of the unit are qualitatively the same as the BVP model.
$g_0(x)$ should be a higher order odd function to make a deep well potential. The negative resistance region on x is localized if the oder of $g_1(x)$ is lower than that of $g_0(x)$. The region is symmetric with respect to $x = 0$ when $g_1(x)$ is an odd function which is the lowest order for the BVP model. If $g_1(x)$ is an even function, the negative resistance region is on an only positive or an only negative side with respect to x. For an example, the region is $0 < x < 1$ for $g(x) = x^3/3 - x^2/2$, and the stable equilibrium point is $x = 1.5$.

4. Stimulus Response

Let us inquire into the stimulus response of the single unit represented by Eq. (5). In the simulation we use $g(x) = (1/\beta)arctanh(x) - x^2, \beta = 3.0, \tau = 1.0, \tau_x = 0.01, W = -10.0, \theta = 0$, and hence, the unit is in a resting state without the stimuli. When θ is a pulse train, the unit fires and generates a single pulse by every single stimulus with a long interval. With decreasing the stimulus pulse interval, the pulse interval ratio between the stimulus pulse train and the firing of the unit decreases discontinuously from 1/1 (single fire by every single stimulus) to 1/2 (single fire by every two stimuli), and from 1/2 to 1/3. Figure 1 shows an example of the unit output and its trajectory on the u - x plane when the ratio is 1/1. We can also observe a chaos-like behavior between the ratio of 1/2 and 1/3. It is shown in Fig. 2.

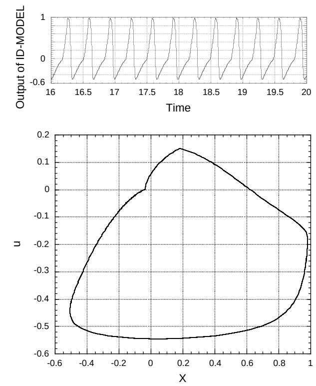

Figure 1: an example of a unit output and its trajectory on the u - x plane

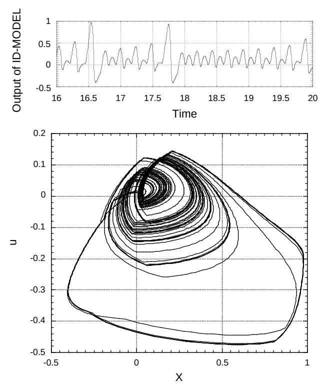

Figure 2: an example of a chaos-like behavior

These behaviors are interesting for analyses of the collective motion of plural units and the network functions.

5. Energy Function

We can discuss the network collective motion by using energy functions. Let us introduce a energy function to the ID model according to Hopfield [6]. However, additional terms are required because of the time delay in Eq. (4) [13]. The energy function is defined as

$$E = -\frac{1}{2\tau}\sum_i\sum_j W_{ij}x_ix_j + \frac{1}{\tau}\sum_i \Theta_i x_i + \frac{1}{\tau}\sum_i \int^{x_i} g_i(x)dx + \frac{\tau_x}{2}\sum_i\left(\frac{dx_i}{dt}\right)^2 . \tag{15}$$

Its time derivative in the case of a symmetric connection strength W_{ij} is

$$
\begin{aligned}
\frac{dE}{dt} &= -\sum_i \frac{dx_i}{dt}\left\{\frac{1}{\tau}\sum_j W_{ij}x_j + \frac{1}{\tau}\Theta_i \right.\\
&\quad \left. -\frac{1}{\tau}g_i(x_i) - \tau_x\frac{d^2x_i}{dt^2}\right\}\\
&= -\sum_i \frac{dx_i}{dt}\left\{\frac{du_i}{dt} + \frac{u_i}{\tau}\right.\\
&\quad \left. +\frac{\tau_x}{\tau}\frac{dx_i}{dt} - \frac{u_i}{\tau} - \frac{du_i}{dt} + \frac{dg_i(x_i)}{dx_i}\frac{dx_i}{dt}\right\}\\
&= -\sum_i \left(\frac{dg_i(x_i)}{dx_i} + \frac{\tau_x}{\tau}\right)\left(\frac{dx_i}{dt}\right)^2. \qquad (16)
\end{aligned}
$$

The coefficient of the right hand side of Eq. (16) is the same as η in Eq. (7). Therefore, the time evolution of the system is a motion in state space that seeks out minima in E and comes to a stop at such points, if η is positive. However, in the negative η region, E turns to increase with time, and the motion climbs the potential hill. It might be expected to get out from local minima by climbing up the hill, if the local minima are negative resistance regions.

A network to answer an optimization problem can be constructed as all units in the network should be at the corners of the solution space when its motion comes to a stop at the global minima. And hence, when it comes to a stop at the local minima, some of the units in the network are in the inside of the space. The negative resistance regions are not at the corners but in the inside of the space when the order of $g_1(x)$ is lower than that of $g_0(x)$ as discussed in the previous section. Therefore, it is expected to increase the possibility to get out from the local minima, if the network is an ID model. Of course, the general way to decide the extent of the negative resistance is not clear at present, it seems to depend on specific problems. However, ID model can provide the intrinsic ability to climb up the potential hill without additional systems like external noises, and hence, it shows one of the solutions to obtain optimum answers. Our preliminary simulations for 4-queen problem always succeed in escaping from local minima and in reaching the global minima.

The behaviours of Hopfield networks strongly depend on their initial values because of the motions on potential surfaces. On the other hand, the utilization of negative resistance decreases the initial value dependency because of the active motions even on potential surfaces.

6. Conclusions

We present the novel model (Inverse Delayed Model) to represent networks with neurons in the critical state, and discuss the fundamental characteristics of the ID model including stimulus responses and an energy function of the the network. The ID model can provide the intrinsic ability to climb up the potential hill without additional systems like external noises, and hence, it shows one of the solutions to obtain optimum answers. It has a simple structure, and hence, it can clear up the relation between the macroscopic network models and the microscopic neuron models.

References

[1] A. L. Hodgkin and A. F. Huxley, " A quantitative description of membrane current and its application to conduction and excitation in nerve, " J. Physiol., 117, pp.500–544, Aug. 1952.

[2] R. FitzHugh, " Impulses and physiological states in thoretical models of nerve membrane," Biophysical J. , 2, pp.445-466, July, 1961.

[3] W. S. McCulloch and W. Pitts, " A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophysics., 5, pp.115–133, 1943.

[4] E. R. Caianiello, " Outline of a theory of thought processes and thinking machine," J. Theor. Biol., 1, pp.204–235, 1961.

[5] M. A. Cohen and S. Grossberg, " Absolute stability of global pattern formation and parallel memory strage by competitive neural networks, " IEEE Trans. System, Man, and Cybernetics, SMC-13, pp.815–826, Sept./Oct. 1983.

[6] J. J. Hopfield, " Neurons with graded response have collective computational properties like those of two-state neurons," Proc. Natl. Sci. USA., 81, pp.3088–3092, 1984.

[7] K. Aihara, T. Takabe, and M. Toyoda, " Chaotic neural networks," Phys. Lett. A., 144, pp.333–340, 1990.

[8] J. Hertz, A. Krogh, and R. G. Palmer, Introduction to the theory of neural computation, Addison-Wesley Publishing Company, Massachusetts, 1995.

[9] T. Nakaguchi, K. Jinno, and M. Tanaka, " Hysteresis neural networks for N-queens problems," IEICE Trans. Fundamentals, E82-A, pp.1851–1859, 1999.

[10] H. Yanai and Y. Sawada, " Associative memory network composed of neurons with hysteretic property," Neural Networks, 3, pp.223–228, 1990.

[11] Krauth, W., Mezard, M., and Nadal, J.-P., " Basins of attraction in a perceptron like neural network," Complex Systems, 2, pp. 387–408, 1988.

[12] J. Nagumo, S. Arimoto, and S. Yoshizawa, " An active pulse transmission line simulating nerve axon," Proc. IRE., 50, pp.2061–2070, 1962.

[13] H. Yanai and Y. Sawada, " Integrator neurons for analog neural networks," IEEE Trans. Circuits Syst., 37, pp.854–856, 1990.