

Generalized McCulloch-Pitts Neuron Model with Threshold Dynamics

Harold Szu and George Rogers,*
NSWC, Code R44, White Oak MD 20903; and *Code K12, Dahlgren VA 22448

Abstract

The McCulloch-Pitts model for a neuron is generalized to endow the axon threshold with a time-dependent nonlinear dynamics. Two components of the threshold vector can be used to generate a pulsed coding output with the same qualitative characteristics as real axon hillocks, that is useful for communication pulse coding.

1. Introduction

The tradeoff consideration between the communication complexity among neurons and the computational complexity of a single neuron has been an important issue in designs of artificial neural networks. McCulloch and Pitts (M-P) proposed their classic model of an artificial neuron in 1943[1]. Advances in neuroscience since the M-P model was introduced have resulted in a much more detailed understanding of biological neurons[2]. Advances in computational capability for artificial neuron model simulation have increased by many orders of magnitude. These two factors make consideration of more detailed models feasible both in terms of plausibility and simulation capability. Recent advances in analog VLSI have made possible analog implementation of some neural network models, but only of very limited size due to the two-dimensional planar implementations and the $O(N^2)$ interconnections required for N neurons in many of the most popular paradigms. Current technology limits a fully interconnected network to sizes on the order of one hundred neurons. This current limitation makes a smarter neuron desirable[3]. In other words, for a fixed number of neurons and interconnections, a path toward enhanced computational capability involves enhancing the computational capability of the individual neurons. For example, a hairy neuron model was proposed by Szu[4] in 1989 to have the pair correlation function. This allows the hairy neuron to prune and grow neurites and thereby grow its own interconnection architecture. Here, we propose a simple dynamical neuron model which can include internal dynamics involving multiple internal degrees of freedom. The proposed model reduces to the M-P model for static inputs and no internal dynamical degrees of freedom. In this communication, we are restricting the treatment to a single neuron without learning. Included are two examples. The ramifications of such neuron models, the treatment of the fully interconnected network case, and more sophisticated applications such as chaotic dynamics will be presented elsewhere [5].

The M-P neuron satisfies the sigmoidal transfer function:

$$v_i = \sigma(u_i) = 1/(1 + \exp(-u_i)). \quad (1)$$

The total input u_i is the net collected from the i th dendritic tree having synaptic junctions indexed by j (conductance denoted by weight w_{ij}), minus a constant threshold θ_i ,

$$u_i = \sum w_{ij}v_j - \theta_i. \quad (2)$$

to be generalized in this paper as a time-dependent vector $\theta_i(t)$ associated with several internal degrees of dynamic freedom. We can obtain two equivalent dynamic versions of the M-P neuron defined by Eqs(1,2). Transposing the left hand side of Eq(2) to the righthand side, $u_i - [\sum w_{ij}v_j - \theta_i] = 0$ as the fixed zero point, and then multiplying both hand sides with an arbitrary time constant

$\alpha\{\text{righthand side}\} = \alpha 0 = 0$, we can assume a local gradient descent that the change of u_i vanishes at the zero point,

$$du_i/dt = -\alpha \{u_i - [\sum w_{ij} v_j - \theta_i]\} = -\alpha \{u_i - \partial H / \partial v_i\} \quad (3)$$

This input dynamics is popularized by Hopfield [6] in terms of a quadratic energy-H landscape concept, namely a Newtonian dynamics-- the force, $-\partial H / \partial v_i$, drives the acceleration of firing rate, du_i/dt . Likewise, Eq(1) is transformed and becomes the fixed point attractor solution of the output dynamics,

$$dv_i/dt = -\alpha \{v_i - \sigma(\sum w_{ij} v_j - \theta_i)\}, \quad (4)$$

as studied extensively by Cohen and Grossberg[7] and recently by Pineda[8]. The two approaches become identical when interconnect w_{ij} and threshold θ_i are time-independent for a top-down design of a fixed neural net architecture. However, when a dynamic threshold $\theta_i(t)$ is needed, we have found that the output dynamics, Eq(4), is more convenient than the input dynamics, Eq(3).

2. Time-Dependent Threshold Dynamics

In order to generalize the M-P neuron model so that a wider range of biological behavior can be modeled, the simplest course of action is to treat the threshold θ_i as a function of various stimuli including the input firing rate u_i . This can be justified on biological grounds. A neuron with an axon hillock structure integrates the input signals to the neuron, and if the integrated potential exceeds a critical threshold, generates pulses, followed by a short refractory period during which no pulses are generated. If the threshold is reached, but the input immediately stops, the axon hillock fires anyway. The magnitude of the input signal governs the rate at which pulses are generated. If the input is of insufficient magnitude for the integrated potential to reach the critical threshold, no pulses are generated.

The dynamical variables in an actual cell that are most important to this pulse generation phenomena are the sodium and potassium currents[9], or equivalently, the populations of sodium and potassium channels. Mead[10] has described an analog VLSI design that behaves in the same qualitative manner as a biological axon. Two internal degrees of freedom allowed us to model an axon hillock. In a real axon hillock, the sodium and potassium currents are the two most important dynamical variables responsible for the generation of the action potential. Thus it is no surprise that an artificial neuron requires two internal degrees of freedom in order to generate an action potential. We adopt the output dynamics Eq(4) as a simple fixed point attractor equation of the threshold, rather than the biologically detailed Hodgkin-Huxley equations[9].

We can equivalently view the M-P neuron equation as the fixed point attractor of either the dynamics of the input variable u_i or the output variable v_i . Although the Hopfield equations are regarded as standard for dynamical neural networks, when introducing internal dynamics to the neuron, it is more natural to consider the time evolution of the neurons output as a function of the inputs and the internal variables. Therefore, we adopt Eq(4) with fixed point attractor Eq(1) as the starting point for generalization to include internal dynamical degrees of freedom.

Suppose we have a dynamic threshold denoted by $\theta_i = \theta_i(\sum w_{ij} v_j, t)$. Then a first order nonlinear dynamical equation for θ_i that is similar in form to Eq(4) is

$$d\theta_i/dt = -\beta \theta_i + \gamma g(\sum w_{ij} v_j, \theta_i) \quad (5)$$

where β and γ are biomaterial constants and $g()$ is some level-off function given in Sect.3.

The ratio γ/β may be determined by the fixed point equation (for constant input)

$$\theta_i = (\gamma/\beta) g(\sum_j w_{ij} v_j, \theta_i) = \theta_i^{(0)}. \quad (6)$$

To produce the oscillatory phenomenon, which is necessary for the refractory and replenishment of ionic populations, a second order linear differential equation is usually needed. Here, we prefer a coupled set of first order (nonlinear or piece-wise linear) differential equations, since multiple degrees of freedom may be easily introduced by converting θ_i from a scalar to a bold face vector

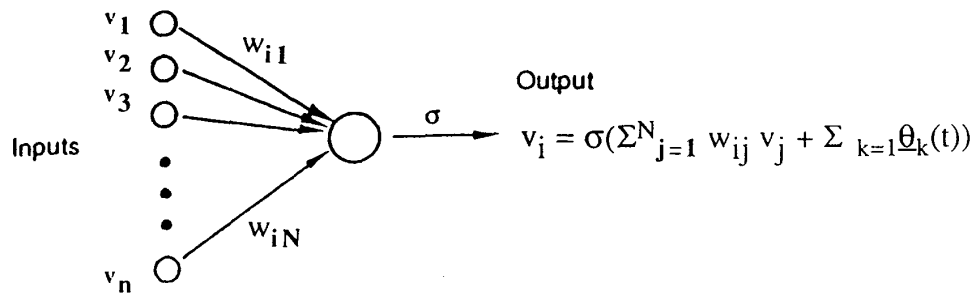
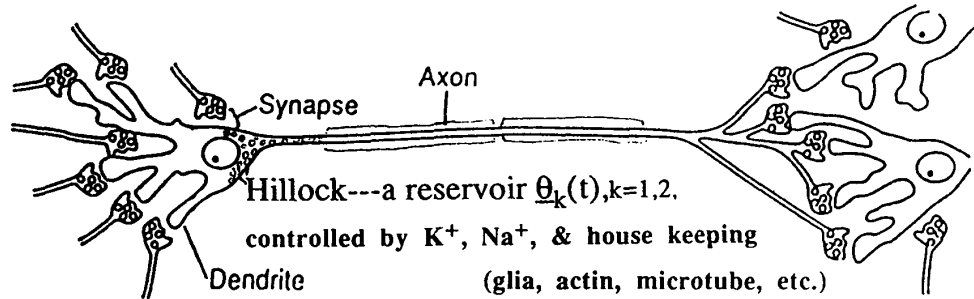
$\theta_{ik}(t)$, with index k indicating components. This results in the following coupled set of equations

$$d\theta_{ik}/dt = -\beta_k \theta_{ik} + \gamma_k g_{ik}(\sum_j w_{ij} v_j, \theta_{i1}, \theta_{i2}, \dots, \theta_{in}) \quad (7)$$

for $k = 1, 2, \dots, n$ where n is the number of internal degrees of freedom. The output dynamics is of the type of a fixed-point attractor:

$$dv_i/dt = -\alpha \{v_i - \sigma(\sum_j w_{ij} v_j - \sum_k \theta_{ik})\}. \quad (8)$$

We have adopted the coupled set of Eq(7) and Eq(8) as our model of medium-grained neuron dynamics, cf. Fig.1 Typical biological and artificial medium-grained neurons.



3. Example

Since we will numerically study a single medium-grained neuron, we will suppress the index i in this section. Usually, a second order differential equation, or two equivalent first order coupled equations, can produce oscillation. By following the biological example and using two internal degrees of freedom, we have developed a simple model based on piecewise linear coupled first order differential equations that duplicates the qualitative behavior of the biological axon. For simplicity, we take the sigmoidal output limit to be the binary threshold logic denoted by the Heaviside function $h(x)=1$ if $x \geq 0$, and zero otherwise,

$$v = h(\theta_1 - 1/2). \quad (10a)$$

The first component of the threshold vector θ_1 is driven by the net input u and a difference threshold

$$d\theta_1/dt = -\alpha_1 \theta_1 + \alpha_1 \sigma(u) + \beta_1 [h(\theta_1 - c) - (3/2) h(\theta_2 - c)], \quad (10b)$$

with respect to the second component equation

$$d\theta_2/dt = -\alpha_2 (\theta_2 - \theta_1), \quad (10c)$$

where $h()$ is the heaviside function, $\sigma(u)$ is given in Eq(1).

With $\alpha_2 = \beta_1 = 0.25 \gg \alpha_1 = 0.01$, this set of coupled equations demonstrates the qualitative behavior of an axon hillock. For a constant input u , the system integrates the input until θ_1 reaches the threshold c . At this point, $\beta_1 h(\theta_1 - c)$ is the dominant term in Eq(10b) and θ_1 will continue to grow exponentially even if u is now set to zero. θ_1 will continue to grow exponentially until $\theta_2 > c$. At this point $-\beta_1 (3/2) h(\theta_2 - c)$ is the dominant term in Eq(10b) causing an exponential decrease in θ_1 . The behavior is largely independent of the other terms because the gain coefficient α_1 is much less than β_1 . Because θ_2 always lags behind θ_1 as can be seen from Eq(10c), before the system recovers from generating a pulse, it will overshoot its resting potential or value. This gives rise to a refractory period after each pulse is generated.

If the input is at such a level that $\sigma(u) < c = 0.25$, the system will never generate pulses. If $\sigma(u) > c$ then the system generates a string of short duration pulses with each pulse followed by a refractory period. The larger $\sigma(u)$ is, the shorter the integration time until the first pulse is generated, and the shorter the period between pulses. The individual pulses that are generated are nearly identical for different input levels. See Fig. 2 for internal threshold components in a close-up time scale, (Cf. Mead[10] p.199, Fig. 12.4), and Fig. 3 the Heaviside neuron output in cases of different sigmoidal output $\sigma(u) = .2, .3, .4, \dots, .9$, and 1.0 due to other neurons' contributions labelled on the coordinate of Fig. 3.

4. Remarks

By considering the M-P neuron as the fixed point attractor of a first order dynamical process, we are led to both the Hopfield and Cohen-Grossberg-Pineda equations. Hopfield Eq(3) treats the input u to the neuron as the dynamic variable while the Cohen-Grossberg-Pineda Eq(4) view treats the output v as the dynamic variable. The Hopfield and Cohen-Grossberg-Pineda equations provide a dual view.

By letting the threshold become a dynamical variable θ subject to its own Cohen-

Grossberg-Pineda like equation, we developed a generalization of the M-P neuron. We then saw that this internal degree of freedom leads naturally to a neuron model which computes a quantity that is sensitive to changes in the input, rather than to the input itself; a neural equivalent of a time derivative. Furthermore, two degrees of freedom in either a second order equation or two first order equations can reproduce the refractory behavior in producing pulse trains. For signal processing, the neural toolbox should include not only the M-P threshold and the fixed-point attractor dynamics, but also a simple neural model of differentiation.

Since the network size is limited by the $O(N^2)$ communication complexity, one can only hope at this time to make each fine-grain processor smarter, medium-grain, and this is our main message. These internal degrees of freedom allow us to numerically reproduce within a single neuron a time pulse train of an arbitrary firing rate. It seems to be appropriate that the community may come forward with VLSI implementations. Thus, the present model is an important contribution toward both neural net pulse coding modeling of communication as well as VLSI design of neural nets.

Acknowledgements

We wish to acknowledge Carey Priebe and Jeffrey Solka for many helpful discussions.

References

- [1]. W. S. McCulloch and W. Pitts, "A Logical Calculus of the Ideas Imminent in Nervous Activity," *Bulletin of Mathematical Biology* Vol. 5, pp. 115-133, 1943.
- [2]. "Frontier in Neuroscience," *Science*, Vol. 242, Nov. 1988.
- [3]. H. Szu, "Neural Networks Based on Peano Curves and Hairy Neurons," *Telemetrics and Informatics*, Vol. 7, pp. 403-430, 1990 (Pergamon Press).
- [4]. H. Szu, "Reconfigurable Neural Nets by Energy Convergence Learning Principles based on extended McCulloch-Pitts Neurons and Synapses," *Int. Joint Conf. neural Networks, IJCNN-89*, p. I-485, Washington D.C. June 1989.
- [5] H. Szu, and G. Rogers, "Single Neuron Chaos," (submitted IJCNN-92)
- [6]. J. J. Hopfield, *Proc. Natl. Acad. Sci. USA*, Vol. 81, pp. 3088-3092, 1984.
- [7] M. Cohen & S. Grossberg, "Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks," *IEEE Trans SMC-13*, 815-826, 1983.
- [8]. F. J. Pineda, "Generalization of Back Propagation to Recurrent Neural Networks," *Phys. Rev. Lett.*, Vol. 59, p. 2229, 1987.
- [9]. A. L. Hodgkin and A. F. Huxley, "Current carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*," *Journal of Physiology*, Vol 116, p. 449, 1952.
- [10]. C. Mead, "Analog VLSI and Neural Systems," Addison-Wesley, 1989 (p.199, Fig.12.4).

Fig. 2 The time evolution of the pulse generation neuron model internal variables during a double pulse integration and refractory period. The first variable θ_1 is shown as a continuous line while the second variable θ_2 is shown as a dashed line. θ_2 always lags behind θ_1 .

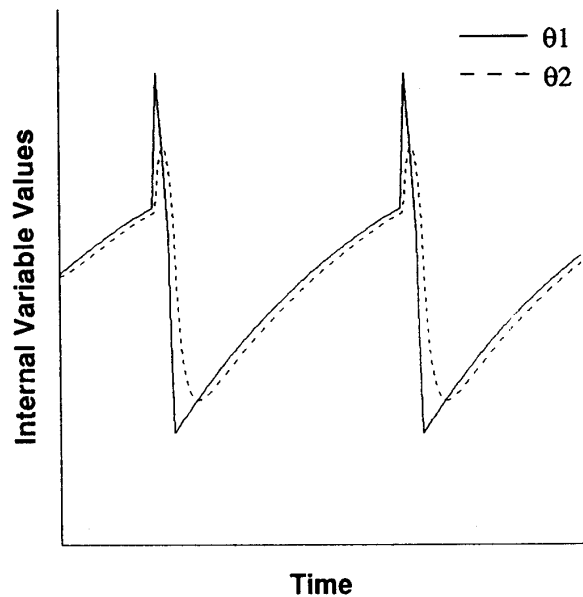


Fig. 3 The output pulse trains of the pulse generating neuron model for different constant input values. As the input level to the neuron increases, the pulse frequency also increases. The actual pulse width, as well as the length of the refractory periods can be adjusted by varying the parameters involved in the equations for v , θ_1 , and θ_2 (Eq(10a-c)).

