Adaptive Gateway Element Based on a Recurrent Neurodynamical Model

Yury S. Prostov* and Yury V. Tiumentsev

Moscow Aviation Institute (National Research University), Moscow, Russia,
*prostov.yury@yandex.ru

Abstract. Dynamic model of a recurrent neuron with a sigmoidal activation function is considered. It is shown that with the presence of a modulation parameter its activation characteristic (dependence between input pattern and output signal) varies from a smooth sigmoid-like function to the form of a quasi-rectangular hysteresis loop. We demonstrate how a gateway element can be build using a structure with two recognizing neurons and one output neuron. It is shown how its functional properties change due to changes in the value of the modulation parameter. Such gateway element can take the output value based on a weighted sum of signals from the recognizing neurons. On the other hand it can perform a complex binary-like calculation with the input patterns. We demonstrate that in this case it can be used as a coincidence detector even for disjoint-in-time patterns. Futhermore, under certain extreme conditions it can be triggered even if only the one input pattern was recognized. Also the results of numerical simulations presented and some directions for futher development suggested.

Keywords: neurodynamic model, hystersis, adaptiveness

1 Introduction

There are a number of challenges in the field of development of an intelligent control systems for highly autonomous robotic systems such as unmanned aerial vehicles. One of these challenges is related to a mechanism that provides the control system with the ability to accumulate the experience from processed data and apply it in further. In other words, it is the task of developing a model with unsupervised or semi-supervised online learning algorithm which can work effectively in a dynamic and uncertain environment under a condition of limited computing resources.

Related and similar tasks are already successfully solved by methods from the field of machine learning such as incremental learning models [1, 2]. They can be applied to an environment in which there are some types of uncertainty and noise or other difficulties. Almost all of them are based on feedforward and simple recurrent architectures which means that they can not reliably maintain activity of neurons without an external signal. But from our point of view this is necessary for context dependent recognition and learning [3]. The LSTM model [4] is the most suitable in this case because it has a special context cell to maintain internal neural network activity. But learning algorithms of this model are based on supervised techniques which is unacceptable in our problem.

We previously proposed the concept of a neural network model [3] as one of the possible approaches to solving the problem outlined above. Our model divided into two parts: working network that performes pattern recognition and learning as well as auxiliary network that evaluates the first one and produces the value of the modulation parameter. An important feature of neuron model used in a working network is that its activation characteristic (i.e. the dependence between an input pattern and an output value) can contain a hysteresis loop [5] under certain values of the modulation parameter as described in [6, 7]. And as shown in [8, 9] the presence of a hysteresis loop in an activation function is related to robust implementations of some working memory models. But at the same time activation characteristic will have the form of a smooth curve under other values of the modulation parameter. Thus, we can change the behaviour of neurons from gradual to trigger mode and can use this property to implement a gateway element with some interesting features.

2 Neuron model

A neuron model used in this paper differs in some details from the one which was described in the related article [7]. Namely, in this article the activation function was replaced by a sigmoidal function and the threshold parameter was moved into a weights vector as one of the coefficients. As a result, model became as follows:

$$\begin{cases} du/dt & = \alpha y + i(\mathbf{w}, \mathbf{x}) - \mu u, \\ y & = f(h(u, \theta)), \end{cases} \tag{1}$$

where $u \in \Re$ is a potential variable, $y \in [0; 1]$ is an output variable, $\mathbf{w} \in \Re^M$ is a weights vector, $\mathbf{x} \in [0; 1]^N$ is an input vector, $\alpha \in [0; +\infty)$ is a recurrent connection weight, $\mu \in (0; 1)$ is a potential dissipation parameter, $\theta \in (0; +\infty]$ is a modulation parameter, $i(\mathbf{w}, \mathbf{x})$ is an external excitation function (in the following we will omit the arguments for brevity) which can be specified as a scalar product or as a Gaussian radial basis function or as any other distance measure function, $h(u, \theta) = u/\theta$ is a potential modulation function, $f(z) = \sigma(z - \Delta)$ is a sigmoidal activation function with $\Delta = 3.0$. Also we assume that the values of parameters α and μ are fixed and selected in advance while the value of a modulation parameter θ is changeable during model operating.

It can be shown that the value of variable y will converge exponentially to some stable equilibrium point y^* of the dynamic system (1). To find these points we need to rewrite equations (1) as follows:

$$F(y) = du/dt = \alpha y + i - \mu\theta g(y) \tag{2}$$

where $g(y) = \Delta + \log(y/(y))$ is the function inverse to the function f. In this case the equilibrium points can be determined from the condition $F(y^*) = 0$.

Moreover, equilibrium point y_j^* will be stable if $F'(y_j^*) < 0$ and unstable if $F'(y_j^*) > 0$, otherwise, additional analysis will be required. It should be noted that there is no analytical solution and therefore this equation must be solved either graphically as shown on Fig. 1 or using numerical methods.

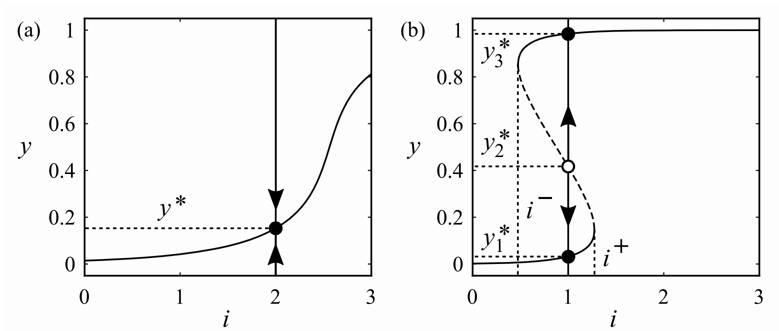

Fig. 1. Graphical solutions of a model (1): (a) the stable point y^* in the case of monostability; (b) the stable points $y_{1,3}^*$ and the unstable point y_2^* in the case of bistability

At the same time bifurcation analysis can be performed analytically based on the Eq. (2). It can be shown that a model (1) has cusp catastrophe [10] by parameters θ and i. Corresponding pitchfork bifurcation at the point $\theta = \alpha/4\mu$ shown in Fig. 2a where the values of parameter i at each point was choosen to get a symmetrical curve. In the case of $\theta \geq \alpha/4\mu$ there exist only one stable equilibrium point and activation characteristic function has the form of a sigmoidal curve as shown in Fig. 1a. Moreover, the slope of this curve decreases as the value of the parameter θ increases.

In the case of $\theta < \alpha/4\mu$ a bistability region arises and it corresponds to the range $i \in (i^-; i^+)$ as shown in Fig. 1b. As we can see increasing of the parameter i value leads to abrupt change of the output value y at the point i^+ and a similar abrupt change occurs at the point i^- during its decreasing. It can be shown that these threshold values $i^\pm$ are determined by the extremes of the equation (2) and can be evaluated as follows: $i^\pm = -\alpha y^\pm + \mu\theta g(y^\pm)$ where $y^\pm = 0.5 \mp \sqrt{0.25 - \mu\theta/\alpha}$. Fig. 2b shows the dependence between these thresholds and modulation parameter θ. As shown in Fig. 2c the values $y^\pm$ themselves determine the region $(y^-; y^+)$ where the stable equilibrium points y^* can not exist. As a result, activation characteristic function takes the form of a hysteresis curve with a loop which becomes closer and closer to a rectangular shape with decreasing value of the modulation parameter θ.

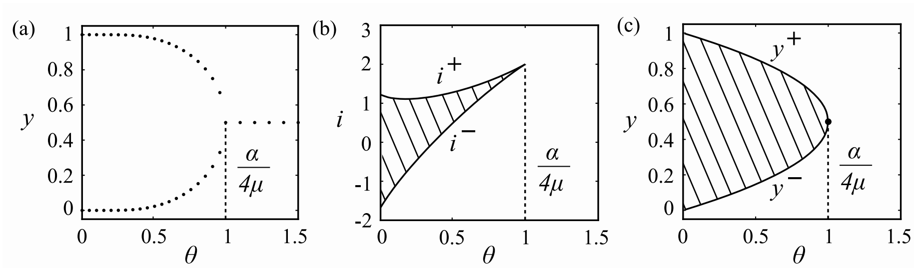

Fig. 2. The properties of the dynamic system (1): (a) a pitchfork bifurcation at the point $\theta = \alpha/4\mu$; (b) the dependence of the thresholds $i^{\pm}$ on the modulation parameter θ (shaded region corresponds to a bistability area); (c) the boundary between areas of a stable (not shaded region) and unstable (shaded ragion) points

3 Gateway model

Let us consider now a gateway model formed by connecting the neurons as shown in Fig. 3a. As a result, overall gateway state will be described as follows:

$$\begin{cases} du_k/dt &= \alpha y_k + i_k - \mu u_k, \\ y_k &= f(h(u_k, \theta)), \end{cases} \tag{3}$$

where the first and second ($k = 1, 2$) neurons process input patterns from two different data channels with appropriate external excitation values $i_{1,2}$ and the third one ($k = 3$) generates the gateway output signal $o = y_3$ using excitation value $i_3 = \beta y_1 + \beta y_2$.

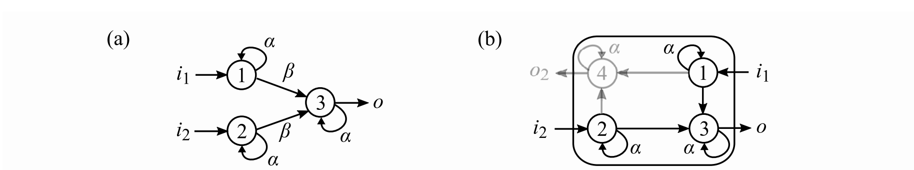

Fig. 3. Architecture: (a) gateway element; (b) computational element for future research based on the model of gateway element

Consider the case when the value of parameter θ corresponds to the region of monostability. As we noted early, in this case activation characteristic of a neuron has the form of a sigmoidal curve. Denote it as a function $\phi_\theta(i)$ where the subscript emphasizes the dependence between the slope of this curve and the value of modulation parameter θ. Then the value of gateway output signal can be represented as $o = \phi_\theta(i_3)$ where $i_3 = \beta \cdot (\phi_\theta(i_1) + \phi_\theta(i_2))$. Thus, the external signals i_1 and i_2 will be transfered by the gateway element in the form of nonlinear weighted sum and the amplitudes of these nonlinear transformations are controlled by the value of modulation parameter θ.

In the case of bistability the value of neuron output variable y takes values from the neighborhoods of points 0 (inactive state) and 1 (active state) as shown in Fig. 2c, i.e. $y \in O^+_\theta(0)$ and $y \in O^-_\theta(1)$ where the subscript emphasizes the dependence between the width of neighborhoods and the value of modulation parameter θ. In this case we can conclude that the value of $i_3 \in O^-_\theta(2\beta)$ if both of values i_1 and i_2 overcome the threshold value i^+ and $i_3 \in O_\theta(\beta)$ if only the one of them overcome the threshold value and otherwise $i_3 \in O^+_\theta(0)$. But the gateway output o can be in active state only when the value of i_3 overcome the threshold i^+. So, for a some fixed range of modulation parameter θ values we can choose the value of parameter β that satisfy to inequality $z_1 < i^+ < z_2$ for $\forall z_1 \in O_\theta(\beta)$ and $\forall z_2 \in O_\theta(2\beta)$. In this case the gateway output o will be active only when both input patterns from data channels are recognized. In other words, the gateway element will become a coincidence detector. But on the other side, we also can choose the value of parameter β which will admit an activation of the gateway element even if input pattern from only the one channel recognized.

Also note the extreme condition when the value of threshold parameter i^- falls below zero as shown in Fig. 2b. In this case the neurons that had previously passed into the active state can remain active even if there is no input signal. As a result, the gateway element can determine a coincidence by disjoint-in-time patterns due to self-sustained activity of recognizing neurons.

We performed numerical simulation to confirm the results obtained above with the following parameters: $\mu = 0.75$, $\alpha = 3.0$ and $\beta = 0.5$. During the simulation we explicitly changed the values of external excitation signals $i_{1,2}$ as well as the value of modulation parameter θ. As shown in Fig. 4 the results of simulation meet with our expectations. The case of performing a nonlinear weighted summation corresponds to the time interval $[t_1; t_2]$. The case of input patterns coincidence detection corresponds to the time interval $[t_3; t_4]$ and the special case of coincidence detection for disjoint-in-time patterns corresponds to the time interval $[t_5; t_6]$.

4 Conclusions

We demonstrated that activation characteristic of the described neurodynamical model of neuron can vary from a smooth sigmoid-like function to the form of a quasi-rectangular hysteresis loop. It was shown how these changes are controlled by the value of modulation parameter θ and how this parameter is related with other parameters of the model.

Also we demonstrated how a gateway element can be build using the described neuron model. It was shown that for the certain range of modulation parameter values the gateway element transfers the input signals as a nonlinear weighted sum but for the other range of modulation parameter values it be-

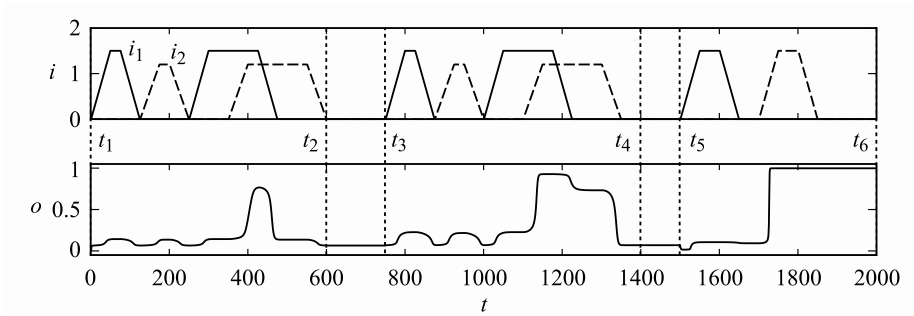

Fig. 4. Results of performed numerical simulation with different values of the external excitation signals i_2 and different values of the modulation parameter θ

gins to perform a binary-like calculation as a complex coincidence detector with additional functional features.

As shown in Fig. 3b futher research is related to the development of a computational element that would learn to associate the ascending and descending data streams in neural network based on the results obtained here for the model of gateway element.

References

1. G. Ditzler, M. Roveri, C. Alippi, R. Polikar, IEEE Computational Intelligence Magazine **10**(4), 12 (2015). DOI 10.1109/MCI.2015.2471196
2. A. Gepperth, B. Hammer, in *European Symposium on Artificial Neural Networks (ESANN)* (2016)
3. Y.S. Prostov, Y.V. Tiumentsev, in *Abstracts of the 12th Intern. Conf. "Aviation and Cosmonautics - 2013"* (MAI (NRU), Moscow, 2013), pp. 619–620
4. S. Hochreiter, J. Schmidhuber, Neural computation **9**(8), 1735 (1997). DOI 10.1162/neco.1997.9.8.1735
5. M.A. Krasnosel'skii, A.V. Pokrovskii, *Systems with hysteresis* (Springer Science & Business Media, 2012)
6. Y.S. Prostov, Y.V. Tiumentsev, in *Proceedings of XVII All-Russian Scientific Engineering and Technical Conference "Neuroinformatics-2015"*, vol. 1 (NRNU MEPhI, Moscow, 2015), vol. 1, pp. 116–126
7. Y.S. Prostov, Y.V. Tiumentsev, Optical Memory and Neural Networks **24**(2), 116 (2015). DOI 10.3103/S1060992X15020113
8. A.A. Koulakov, S. Raghavachari, A. Kepecs, J.E. Lisman, Nature neuroscience **5**(8), 775 (2002). DOI 10.1038/nn893
9. C.D. Brody, R. Romo, A. Kepecs, Current opinion in neurobiology **13**(2), 204 (2003). DOI 10.1016/S0959-4388(03)00050-3
10. E.C. Zeeman, *Catastrophe theory: Selected papers, 1972–1977.* (Addison-Wesley, 1977)