Spiking Neural Network with Lateral Inhibition for Reward-Based Associative Learning

Nooraini Yusoff and Farzana Kabir Ahmad

School of Computing, College of Arts and Sciences,
Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
nooraini@uum.edu.my

Abstract. In this paper we propose a lateral inhibitory spiking neural network for reward-based associative learning with correlation in spike patterns for conflicting responses. The network has random and sparse connectivity, and we introduce a lateral inhibition via an anatomical constraint and synapse reinforcement. The spiking dynamic follows the properties of Izhikevich spiking model. The learning involves association of a delayed stimulus pair to a response using reward modulated spike-time dependent plasticity (STDP). The proposed learning scheme has improved our initial work by allowing learning in a more dynamic and competitive environment.

Keywords: Lateral inhibition, Spiking neural network, Associative Learning, Spike-time dependent plasticity.

1 Introduction

It has been evidently known that, in many parts of the brain, networks are recurrent in nature with sparse connectivity, e.g., [6],[7]. In the systems with sparse representation, neurons cooperate and compete with each other to accomplish a task. It has also been proposed that lateral inhibition plays a key role in many of the brain's fundamental computational abilities. Nevertheless, the underlying mechanism in a neural system with such sparse representation still remains intriguing. In a dynamic and competitive environment, not much is known how a lateral inhibition acts as a filtering apparatus in information processing to provide more intense representation of stimuli.

In this study, we show how a lateral inhibition between neuronal groups can be solved via synapse reinforcement based on reward modulated learning. Given a learning setting with some degrees of correlation in spike patterns, during a response interval time, the proposed learning scheme first triggers the network inhibitory response groups to depress activations of their competitors, and then strengthens the connectivity to its target excitatory response groups. The reinforcement signal is dependent on activation rate (i.e. firing activity) in response groups. The lateral inhibition results in stronger synapses in both target inhibitory and excitatory pathways.

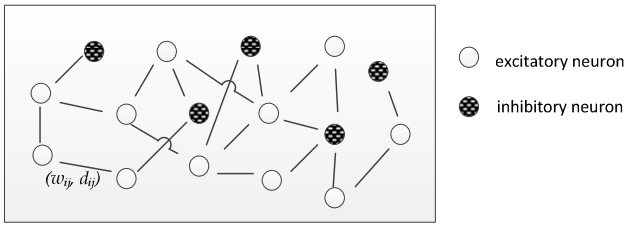
C.K. Loo et al. (Eds.): ICONIP 2014, Part I, LNCS 8834, pp. 327–334, 2014.

1.1 Initial Work

From our preliminary work [3], we introduced a pair-associate learning for stimulus-stimulus-response (S-S-R) association. The learning scheme trains a spiking neural network to associate a delayed stimulus pair to a response. The first stimulus is presented to the network, followed by the second stimulus after a delay, the activity of the response subpopulations is then observed within an interval. The response group with highest activation rate is considered to be the winner.

The simulation model was a spiking neural network with random and sparse connectivity (probability p=0.1) consisting of 1000 neurons (80% of excitatory and 20% inhibitory neurons). The network has random synaptic transmission delays between 1 to 20 ms [5],[8]. The spiking dynamics of a neuron follow the properties of Izhikevich model [4]. The excitatory synapses are plastic whilst, the inhibitory synapses are not plastic.The excitatory neurons population is divided into subpopulations of m stimulus groups S, n response groups R and non-selective neurons NS. In the initial model, the inhibitory subpopulation IH acts as global inhibition (Fig. 1).

A.



B.

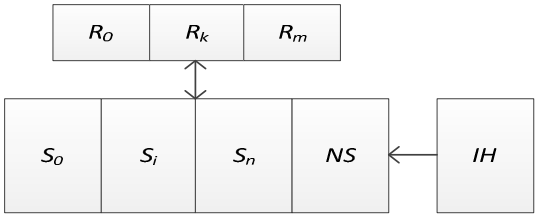


Fig. 1. (A) Schematic view of a recurrent spiking neural network consisting of 80% excitatory (N_E) neurons and 20% of inhibitory (N_I) neurons, with sparse and random connectivity, p = 0.1 (no self-feedback), i.e. $N_E \rightarrow \{N_E, N_I\}$ and $N_I \rightarrow N_E$. Each synaptic transmission has random delay $d \in [1, 20]$. (B) Neurons are divided into subpopulations of stimulus groups (S), response groups (R), non-selective neurons (NS) and inhibitory pool (IH). S and R are composed of 50, and 100 excitatory neurons, respectively.

With a simple network structure in learning we implemented the winner-take-all (WTA) strategy via application of random excitatory bias signals to the winner of target response groups. With the WTA method, it could increase the probability of activation of some neurons in a target response group that had not been fired. This would conse-

quently result in higher activation in the target response group compared to its competitors. However, the simplicity of the structure has some limitations for learning with high competition. For learning with high correlation in spike patterns, the model performance decreased due to undesired causal firings, e.g. when the network was trained to associate $(S_0, S_1) \rightarrow R_A$ and $(S_0, S_2) \rightarrow R_B$, with two competing responses, i.e. neural subpopulations R_A and R_B. Furthermore, strengthening of synaptic strength between $S_i \rightarrow R_A$ could also lead to activation of neurons in response group R_B due to triggering of synapses $R_A \rightarrow R_B$, i.e. firings of postsynaptic neurons of R_A in R_B.

2 Network with Lateral Inhibition

To improve the discrimination rate in a competitive learning, we suggest a modified network topology with a lateral inhibition mechanism (see Fig. 2). In the network (consisting of 1000 neurons) with lateral inhibition, we eliminate the excitatory synaptic connections between response groups. Excitatory neurons in each response group, e.g. R_{+m},are connected to their inhibitory pool, e.g., R_{-m}. The inhibitory pool provides inhibition to its competitor group(s) through negative synaptic connections.

The synaptic strength from an inhibitory pool of a response group to the excitatory neurons in its competitor is set to -4.0 (a strong inhibition). Generally, each neuron has connectivity of 0.1 (i.e. 100 out of 1000 neurons). Each excitatory neuron in the response groups has 50 postsynaptic neurons from its inhibitory pool, and 50 postsynaptic neurons consisting of neurons from the same excitatory response group and/or excitatory neurons in the input module. Meanwhile each inhibitory neuron in the response groups is connected to other 100 excitatory neurons of its competitor groups. By having such anatomical constraint in the response module, activation of any neuron in a response group will invoke its inhibitory pool that eventually sends out some amount of inhibitory postsynaptic potentials to its competitor(s).

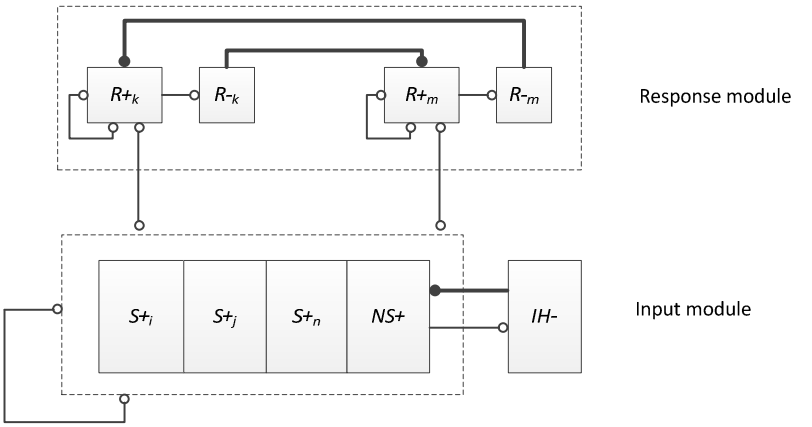


Fig. 2. Recurrent spiking network with subpopulations of stimulus groups (*S*), response groups (*R;*R_+ *and* R_-), non-selective neurons (*NS*) and inhibitory pool (*IH*). Lines end with open circle show excitatory connections, and lines end with solid circle indicate inhibitory connections.

For synaptic connections in stimulus neurons population (i.e. *S* and *NS*), each excitatory neuron has random connections to 100 neurons from the whole populations (from 1000 neurons), and each inhibitory neuron in this module is connected to 100 excitatory neurons from the whole population as in the network without lateral inhibition in our earlier model.

3 Stimulus-Stimulus-Response Associative Learning

All training simulations presented in this paper were implemented in C++ and testing or probe trials were performed in MATLAB.

3.1 Simulation Method

For stimulus representation, we randomly select 50 neurons from each group to deliver a superthreshold current of 20 pA, for example in group S_0 consisting of 100 neurons, 50 neurons are selected to be paired with 50 neurons from group S_1 (out of 100 neurons, chosen randomly). Hence for two stimulus pairs, e.g. $(S_0,S_1) \rightarrow R_A$ and $(S_0,S_2) \rightarrow R_B$, the stimulus S_0 might have a number of overlapping neurons.

In a 20-minute simulated time, we implement an association of a set of stimulus pairs to their target responses. The learning is initialised with a random background activity for 100 ms. During the initialisation phase, we stimulate an arbitrary neuron with 20-pA (strong) current for every ms. With the same random background activity, we present to the network a pair of stimulus (S_i,S_j), selected randomly, via intensification of 1-pulse current (i.e. 20 pA) to all neurons in the selected stimulus groups. After that, a group S_i is stimulated, followed by its associated pair S_j after an interstimulus interval (ISI). An optimal ISI is chosen from a range of 10 – 50 ms based on a preliminary experiment.

From the onset of the second stimulus, we count the number of activations in the response groups, R_k, within 20 mstime interval. The response group with the highest number of activations is considered to be the winner. The next learning pair is presented after a 100-ms delay from the offset of each response interval. The learning result reported in this paper, is an averaged performance of 10 simulated networks.

For a testing phase also known as the "probe trial", we run a simulation consisting of a number of trials for 200 ms each. In each trial, we present a stimulus pair to the network randomly with equal probability for each pair to be tested. We also apply some degree of distortion via smaller random activation of neurons in a learned stimulus group i.e. with probability of less than 1.0.

The network with some background activity (for the first 100 ms in each trial) as described before is then intensified with super threshold current of 20 pA applied onto the tested prime stimulus at some random time, *t* in between 100-120 ms, i.e. after the random activity. The stimulation of its pair group proceeds after the prime stimulus group depending on the tested ISI. The number of spike counts within the 20-ms response interval (starts from the onset of the *choice*) is used to compute a winning response. The testing result expresses the averaged percentage of performance over a number of trials, i.e. performance = (number of correct recall/number of trials)*100.

3.2 Learning Rules

The synaptic efficacy is dependent on the reward signal *r(t)* (2). The signal modulates the synaptic changes read from a spike-timing dependent plasticity (STDP) function (as in 1).

$$\Delta w_{stdp} = \Theta \{ A_+ e^{-\Delta t/\tau +}, \Delta t \geq 0; A_- e^{\Delta t/\tau -}, \Delta t < 0 \} \quad (1)$$

From (1), the synapse is potentiated if the difference in firing times (Δt) between a postsynaptic neuron and its presynaptic neuron (i.e. t_{post}-t_{pre}) is ≥ 0, otherwise the synapse is depreciated. The magnitude of potentiation (depression) is given by $A_+ e^{-\Delta t/\tau +}$ ($A_- e^{\Delta t/\tau -}$), where A represents the maximal change when the spike timing difference Δt is approaching 0, and τ is the time constant (in ms). For our STDP curve, $\tau_+ = \tau_- = 20$ ms, $A_+ = 0.1$, and $A_- = 0.15$ [2].

The reward signal *r(t)* determines the amount of modulation to the summation of Δw_{stdp}. Therefore, the reward modulated STDP learning holds [1], [2]:

$$\Delta w(t) = [\alpha + r(t)]\, z(t) \quad (2)$$

where α is the activity-independent increase of synaptic weight, and *z(t)* represents the summation of Δw_{stdp} obtained from (1). Excitatory and inhibitory weights are initialised to 1.0 and -1.0, respectively. To avoid infinite growth, weights are kept to be in the range between 0 and 4 mV.

3.3 Synapse Reinforcement

Synapse reinforcement is implemented based on a reward policy. The reward policy determines the amount of synapse potentiation (i.e. strong or weak potentiation) or depression. The network is given a strong positive reward, $r(t-1) + 0.5$, if a target response group, e.g. R_A is the winner having neuron firing rate (F) in the groupgreater or equal than twice of its closest competitor, e.g. R_B, or a weak reinforcement signal, $1-(F_{R\_A}/ F_{R\_B})$ if the neuron firing rate is greater than (and less twice of) its closest competitor. Meanwhile, the network receives a negative reward signal -0.1 if $F_{R\_A} < F_{R\_B}$.

Synapse reinforcement is implemented in two phases. In the first phase, within the 20-ms interval, we reward the network based on the number of activations in the response inhibitory groups within the first 10 ms. This is to strengthen the synapses for connectivity between a stimulus and the target response inhibitory group for preventing the activation of response competitor groups. In the second phase, we reward the network for the number of activations in the response excitatory groups within the 20-ms response interval for synapse reinforcement from the stimulus group to the target response excitatory group.

4 Results

4.1 Correlation in Spike Pattern

As discussed in Section 1.1, we initially trained a network with fully sparse and random connectivity. We trained the network with a set of learning pairs consisting of exclusive stimulus neurons groups, *Pair-Response* = $\{(S_0,S_1) \rightarrow R_A, (S_2, S_3) \rightarrow R_B, (S_4, S_5) \rightarrow R_A, (S_6, S_7) \rightarrow R_B\}$. As a result of learning, the average number of spikes for target response groups is 9.98, when compared with the non-reinforced groups with 7.18 and the negatively rewarded groups with 3.15. The correct memory recall was achieved at 99.9%.

We further experimented the learning with non-exclusive stimulus groups to see the effect of spike correlation for three conditions of learning pairs, condition I – shared the first stimulus, *Pair-Response* = $\{(S_0,S_1) \rightarrow RA, (S_0,S_2) \rightarrow B\}$, condition II – interference from non-exclusivity with identical orthogonality, *Pair-Response* = $\{(S_0,S_1) \rightarrow R_A, (S_0,S_2) \rightarrow R_B, (S_1,S_0) \rightarrow R_B\}$, and condition III - non-exclusivity with asymmetrical difference, *Pair-Response* = $\{(S_0,S_1) \rightarrow R_A, (S_0,S_2) \rightarrow R_B, (S_2,S_1) \rightarrow R_A\}$. To create more interference effects due to neural spike train correlation, the ISI was set to 10 ms as the average of synaptic delays in the range of 1 to 20 ms. The results are exhibited in Table 1. From Table 1, the results demonstrate the level of interference that could disrupt the stability of a pattern due to conflicting responses. The effect of non-exclusivity could be observed when any of learning pairs shared the first or second stimulus.

4.2 Learning with Lateral Inhibition

We repeated the learning experiment with non-exclusive groups for the network with lateral inhibition as described in Section 2.In all the three conditions, learning performance could be improved through implementation of our proposed lateral inhibition (see Table 1). For training, the averaged discrimination rates in conditions I, II and III are 86.56%, 76.99% and 93.79%, respectively, in comparison with learning without the lateral inhibition, 53.89%, 46.30% and 78.26% for conditions I, II and III, respectively.

Table 1. Correct memory recall to target response for condition I – shared the first stimulus, condition II – interference from non-exclusivity with identical orthogonality, and condition III - non-exclusivity with asymmetrical difference

Condition	*Correct memory recall (%)*	
	No lateral inhibition	With lateral inhibition
I	50.25	85.40
II	47.33	73.73
III	83.60	96.00

We also ran memory recalls for noisy stimuli with only a fraction of neurons stimulated randomly with $0.7 \leq p \leq 1.0$. The results were as follows for the distorted test

pairs, 70.80%, 70.07% and 81.60% respectively (without lateral inhibition: 51.13%, 44.40% and 70.13%). Furthermore, the network was also trained with non-exclusive stimulus groups with stimulus pairs as follows (multiple responses): *Pair-Response* = $\{(S_0,S_2)\rightarrow R_A,\ (S_0,S_3)\rightarrow R_B,\ (S_1,S_2)\rightarrow R_C,\ (S_1,S_3)\rightarrow R_D\}$. The correct recall rate was achieved at 78.47% and 78.70% for training and testing, respectively.

5 Conclusion

Initially, learning tasks only involved association to two response groups, R_A and R_B. In such cases, neurons in both groups act as the dopamine neurons whose activation within its group in an interval time could be a behavioural action in anticipation of the reward.

There was a limitation due to high correlation of spike patterns that might cause instability of learning pairs. We have analysed several levels of interference that can lead to high competition of responses. Even the performance in some cases was above chance, non-inclusivity in learning pairs could somehow affect discrimination of temporal sequences. For example for a system with shared stimulus, e.g. *Pair-Response* $\in\{(S_0,S_1)\rightarrow R_A,\ (S_2,S_1)\rightarrow R_B\}$, any of the stimulus pairs could be dragged to an undesired response. As an immediate solution, we introduced some anatomical constraints on the current network model by eliminating the excitatory connections and inserting inhibitory connections between neurons in response groups. This provides a solution to enhance the discrimination rate for some learning conditions with non-exclusive stimulus groups. As learning progresses, reinforcement of synapses is achieved not only to target response groups but also to its inhibitory pool from a triggered stimulus pair. Strengthening of synaptic connections to an inhibitory pool could facilitate discrimination of a target group as neurons in the competitor groups will be suppressed.

We have improvised the excitatory-inhibitory network as proposed in [2] and [11] by adding lateral inhibiton connections that can prevent activations of non-desired responses. Even though the biological interpretation of such an inhibition mechanism is not well defined in our model, this serves as an initial attempt for understanding the synapses of the anterior cingulate cortex (ACC) triggered on events related with conflict or error detection, e.g., [9], [10].

We have also tested learning in environments with higher competition of responses. We extended the training to discriminate paired stimuli for four responses, i.e. R_A, R_B, R_C and R_D. Moreover, using the real images data, we have also performed learning for visual recognition task. The training result achieved at 89.46% and all image pairs were correctly discerned with 100.00% accuracy in probe trials [12].

The performance indicates some potential of our model in learning multiple input-output mappings with high competition of outputs. Nevertheless, the increase in the number of responses requires greater number of spikes from the input neurons with minimum of 80% activation from each stimulus group.

Acknowledgements. This research has been funded by the Ministry of Higher Education (Malaysia).

References

1. Florian, R.V.: Reinforcement learning through modulation of spike-timing dependent synaptic plasticity. Neural Comput. 6, 1468–1502 (2007)
2. Izhikevich, E.M.: Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb. Cortex 17, 2443–2452 (2007)
3. Yusoff, N., Grüning, A.: Learning Anticipation through Priming in Spatio-temporal Neural Networks. In: Huang, T., Zeng, Z., Li, C., Leung, C.S. (eds.) ICONIP 2012, Part I. LNCS, vol. 7663, pp. 168–175. Springer, Heidelberg (2012)
4. Izhikevich, E.M.: Simple Model of Spiking Neurons. IEEE Trans. Neural Networks 14(6), 1569–1572 (2003)
5. Izhikevich, E.M.: Polychronization: Computation with Spikes. Neural Computatio. 18, 245–282 (2006)
6. Jones, E.G.: Microcolumns in the cerebral cortex. Proceedings of the National Academy of Sciences 97(10), 5019–5021 (2000)
7. Cutsuridis, V., Wennekers, T.: Hippocampus, microcircuits and associative memory. Neural Networks 22, 1120–1128 (2009)
8. Paugam-Moisy, H., Martinez, R., Bengio, S.: Delay learning and polychnization for reservoir computing. Neurocomputing 71(7-9), 1143–1158 (2008)
9. Botvinick, M., Braver, T., Barch, D., Carter, C., Cohen, J.: Conflict monitoring and cognitive control. Psychological Review 108(3), 624–652 (2001)
10. Kaplan, G.B., Sengor, N.S., Gurvit, S., Guzelis, C.: Modelling The Stroop Effect: A Connectionist Approach. Neurocomputing 70, 1414–1423 (2007)
11. Brunel, N., Lavigne, F.: Semantic Priming in a Cortical Network Model. Journal of Cognitive Neuroscience 21(12), 2300–2319 (2009)
12. Yusoff, N., Grüning, A.: Biologically Inspired Temporal Sequence Learning. Journal of Procedia Engineering 41, 319–325 (2012)