Charles University in Prague
First Faculty of Medicine

Dissertation Thesis Resume

Activity and Memory in Biologically Motivated Neural Network

Author: Mgr. Július Štroffek

2018

Doktorské studijní programy v biomedicíně

Univerzita Karlova v Praze a Akademie věd České republiky

Obor:	Biomedicínská informatika
Předseda oborové rady:	prof. MUDr. Štěpán Svačina, MBA, DrSc.
Školící pracoviště:	Ústav patologické fyziologie 1.LF UK
Školitel:	prof. MUDr. RNDr. Petr Maršálek, Ph.D.

Disertační práce bude nejméně pět pracovních dnů před konáním obhajoby zveřejněna k nahlížení veřejnosti v tištěné podobě na oddělení pro vědeckou činnost a zahraniční styky děkanátu 1. lékařské fakulty.

Program of Doctoral Studies in Biomedicine

Charles University in Prague and
The Academy of Sciences of the Czech Republic

Subject:	Biomedicínská informatika
Chair of the subject board:	Prof. MUDr. Štěpán Svačina, MBA, DrSc.
Department:	Institute of Pathological Physiology, First Faculty of Medicine, Charles University in Prague
Advisor:	Prof. Petr Maršálek, MD, Ph.D.

The thesis will be available to public at least five working days prior the defense at the Department of Research and Foreign Relations of the Dean's Office of the First Faculty of Medicine, Charles University in Prague.

Contents

Abstract (in Czech language)

Tato práce prezentuje biologicky motivovaný model neuronové sítě, který funguje jako autoasociativní paměť. Architektura prezentovaného modelu odpovídá architektuře Hopfieldovy sítě, jež může odpovídat některým částem, které byly identifikovány v hipokampální oblasti CA3 (Cornu Amonis). Vzory v modelu nejsou statické stavy neuronů, ale cyklicky se opakující synchronní aktivity s nízkým relativním počtem současně aktivních neuronů. Vzory jsou do sítě uloženy pomocí Hebbova pravidla upraveného na ukládání sekvencí. Navrhnutý model je analyzován z pohledu kapacity spolu s numerickými simulacemi. Model je dále rozšířen o krátkodobé posilování synapsí (STP), které je v modelu nutnou součásí správného vybavování vzorů. Důsledkem tohoto rozšíření je další výrazné zvýšení kapacity modelu. V práci je diskutována možnost kombinace obou přístupů. Síť může zpracovat vzory v krátkém časovém intervalu bez STP (rychlé vzory) nebo pomocí STP v delším časovém intervalu (pomalé vzory). Z vlastní zkušenosti víme, že některé vzory se mohou vybavit rychle a některé k vybavení potřebují daleko delší čas.

Abstract

This work presents biologically motivated neural network model which works as an auto-associative memory. Architecture of the presented model is similar to the architecture of the Hopfield network which might be similar to some parts of the hippocampal network area CA3 (Cornu Amonis). Patterns learned and retrieved are not static but they are periodically repeating sequences of sparse synchronous activities. Patterns were stored to the network using the modified Hebb rule adjusted to store cyclic sequences. Capacity of the model is analyzed together with the numerical simulations. The model is further extended with short term potentiation (STP), which is forming the essential part of the successful pattern recall process. The memory capacity of the extended version of the model is highly increased. The joint version of the model combining both approaches is discussed. The model might be able to retrieve the pattern in short time interval without STP (fast patterns) or in a longer time period utilizing STP (slow patterns). We know from our everyday life that some patterns could be recalled promptly and some may need much longer time to reveal.

Introduction

Memory is an ability of the organism or system to store, retain and recall information. In biology memory is mostly associative and stored pattern is recalled as a response to some other input activity previously perceived or being very close to the activity previously perceived. The stimulus that leads to the response is called input and the response is called output. There are multiple types of associative memories and possibilities how they could be implemented.

Usually, the same memory can store lot of information and during a process of storing and recalling information an addressing of the proper piece has to take place. There are two technical approaches how to address information, which leads to two different types of memories.

Location addressable memory uses the exact location of the information to store and recall the appropriate information. This is the way computer memory works or a sheet of a paper with notes is used. This type of memory is mostly used by computers or other machines. When using a computer memory the location is specified by the memory address.

In *content addressable memory* we store and recall information based on its content. This type of memory is often called *associative memory*. It is natural to most of the organisms and the human brain. The stream of thoughts of our internal monologue (James, 1982) uses the content of thoughts to address the stored information. The thoughts are of course influenced by our perception. The recalled information may have a form of some movement, thought or some chemical reactions. If we feel hunger and smell a good food we would have a clear picture of the food just from its smell even without seeing it. This may lead even to spontaneous salivation without seeing a food as studied by I. P. Pavlov on dogs under different conditions in (Pavlov, 1927). Pavlov described a conditional reflexes and processes that led to strengthening or weakening of those reflexes. From the memory point of view, these reflexes might be just a store and retrieval processes to/from some associative memory.

The view on types of memory is different in physiology and psychology. Memory is classified based on the experience it brings to individual. The memory could be declarative or procedural. The declarative memory is used consciously to store or recall information. It could be further divided into semantic memory containing the abstract information and episodic memory storing the contextual information related to time, place, emotions etc.

Other memory classification divides the memory into sensory memory, short term memory and long term memory. The short term memory is transferred into long-term memory by memory consolidation process. This classification was proposed in (Atkin, 1968). However, all these types of memory might have the same or similar underlying neural mechanisms that form their behavior.

In our thesis we simply understand the information as the real vector or the array of real numbers that correspond to some type of neural activity and we do not take care of any further interpretation. We will focus on associative models of neural networks.

One of the most studied memory models in neural networks are auto-associative memories. In this case, the network does not need any separate input and output neurons. The input is presented to the network as external excitation or simply by setting up the neuron outputs. The network afterwards evolves based on its dynamics and it converges to the stable state. This stable state is considered to be the output of the network. This is achieved by the recurrent connections in the network where output of each neuron can be potentially connected as input to all other neurons.

Artificial neural networks usually distinguish between learning and recall processes. The learning process is used to set up synaptic efficiencies in a way that will allow the model to retrieve the required patterns. This is often based on the Hebb's hypothesis (Hebb, 1949; Kuriščák et al. 2015). During recall process these synaptic efficiencies do not change and the network dynamics is used to recall the corresponding pattern. It was shown in (Tsodyks and Markram 1997; Tsodyks et al. 1998) that the changes in synaptic efficiencies could occur in a time scale small enough to also affect the pattern recall processes.

Goals of the Thesis

Initial motivation for our work were the papers (Tsodyks and Markram 1997; Tsodyks et al. 1998) where the dynamics of detailed neural synapse was presented. It was shown also in biological experiments that the time window for the change of the synaptic efficiency could be as low as few milliseconds having a recovery in range of seconds (Tsodyks et al. 1998). This offers the possibility to include the changes of synaptic weights into recall phase of neural network models.

Our goal was to find some improvement of the existing models that will be extended by this dynamics and will improve the model performance. First, we have build our model on Hodgkin and Huxley equations (Hodgkin and Huxley, 1952) with the synapse implemented according (Tsodyks et al. 1998). We have observed improvements in capacity and further we wanted to simplify our model as much as possible for further analysis and simulations while keeping its improved properties. We ended up with a combination of Willshaw and Hopfield models which we further studied in more detail. Our main results were published in papers which are attached to this thesis.

Materials and Methods

We have defined the model and we have investigated it's memory retrieval properties by computer simulations (Štroffek et al., 2007). We have optimized the learning algorithm for sparse patterns as the number of stored patterns was very large (Rolls and Treves, 1998).

The model we used has binary neuron outputs, binary synaptic weights as in Willshaw model (Willshaw, 1969). We used the network as auto-associative memory with the topology and network dynamics as in Hopfield network (Hopfield, 1982). We have used cyclic patterns that we stored into our network model (Štroffek et al., 2007).

We have altered the Hebb's rule (Štroffek and Maršálek, 2012) to learn the cyclic activities. We have decomposed the cyclic patterns into sub-patterns which correspond to the state of neurons in the specific discrete time step. We have altered the learning rule originally used for Hopfield network to the form required for cyclic pattern storage.

Let us denote the length of i-th pattern as $l(i)$, the successive sub-patterns of the i-th pattern as ${}^{i}t^{1}, {}^{i}t^{2} \ldots {}^{i}t^{l(i)}$ and the activity of n neurons in j-th sub-pattern as ${}^{i}t^{j}_{1}, {}^{i}t^{j}_{2} \ldots {}^{i}t^{j}_{n}$.

Now, let us have p patterns which we would like the network to learn. In each k-th iteration of the learning process the network learns the k-th pattern (a cycle of sub-patterns). Let us denote weights after the k-th iteration as w^{k}_{ij}. In each iteration we modify the weights according to the equation:

$$w^{k}_{ij} = \max \left(w^{k-1}_{ij}, \mathrm{H} \left({}^{k}t^{l(k)}_{i} \; {}^{k}t^{1}_{j} + \sum_{q=1}^{l(k)-1} {}^{k}t^{q}_{i} \; {}^{k}t^{q+1}_{j} \right) \right). \tag{1}$$

The pattern retrieval test could not be efficiently optimized for sparse patterns. Therefore, we made a random selection of stored patterns that we have tested for retrieval. The patterns learned were randomly generated in all tests performed. The algorithm used for pattern generation is described in (Štroffek et al. 2007).

We have used two types of pattern retrieval tests. First test type was focused on testing that the patterns that were stored to the network are

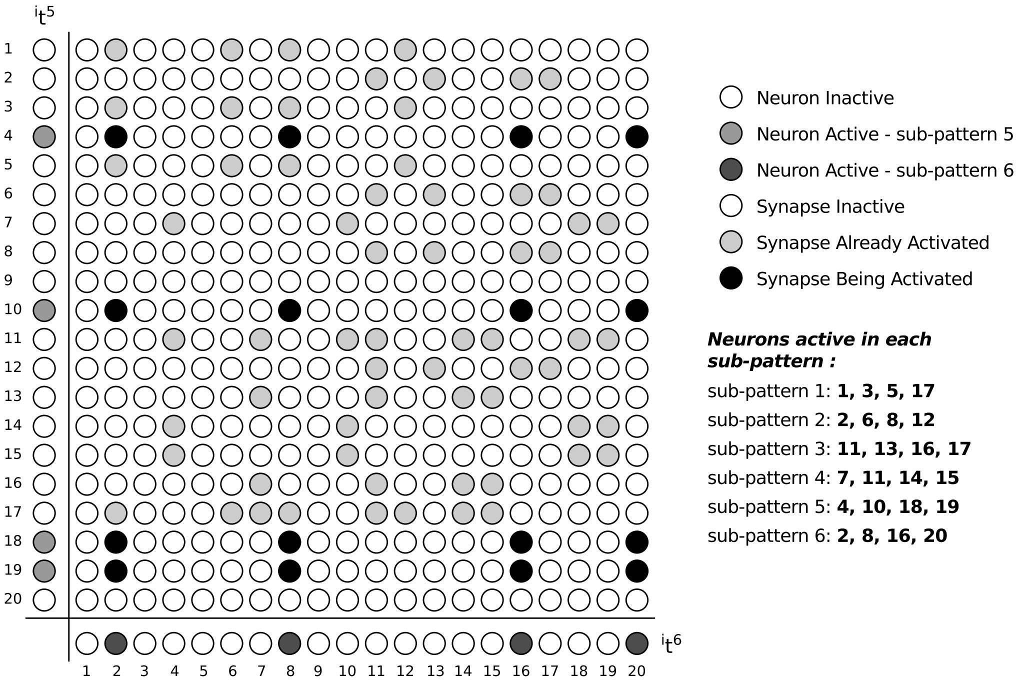

Figure 1: **Schematics of the learning algorithm.** The pattern learned here consists of 6 sub-patterns. This cartoon shows how the synapses are activated based on the neuron output values in example sub-patterns 5 and 6. The left column corresponds to neurons in sub-pattern 5, the bottom row corresponds to neurons in sub-pattern 6.

successfully retrieved. We called this type of test as *positive memory test*. The second type of the test was focused on testing that patterns that were not stored to the network are not successfully retrieved. We called this type of test as *negative memory test*.

Results

We have shown that if we use cyclic patterns and we extend the synapse with short-term potentiation like behavior the capacity of the model could be increased.

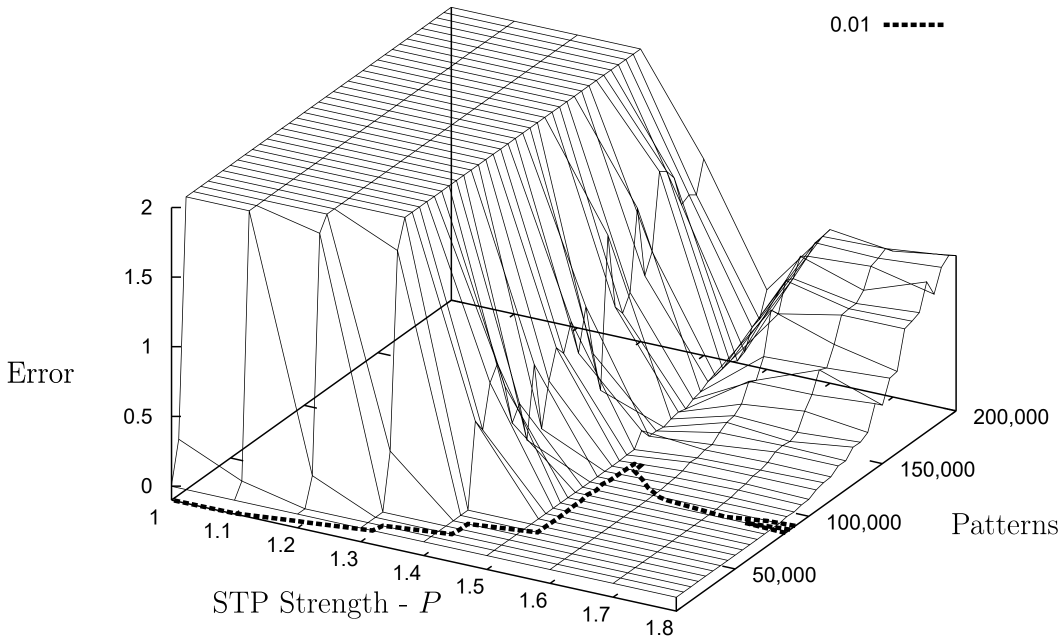

Figure 2: **Capacity obtained by negative memory test.** There is a thick contour line drawn for the relative error of 0.01 (1 %). The local minimum is visible for $P = 1.5$ in the surface and the contour. We show only negative test as it restricts the capacity of the model more than the positive memory test.

It could be seen from the given figure that the memory capacity could be highly increased by STP even 5 to 10 times compared to the same model with no STP.

We made a memory estimate based on the ratio of activated synapses, (Stroffek et al., 2007). Let us assume that the ratio of activated synapses is K. We denote the relative activity in sub-patterns as a and the estimated number of sub-patterns stored as p. Then the estimation is

$$p \approx \frac{(1-K)}{(1-a^2)} \tag{2}$$

Discussion

We have presented very simple neural network as a combination of Willshaw and Hopfield models having binary neuron outputs as well as binary synaptic weights. We have changed the way how patterns are encoded and we additionally have extended the model with simple synaptic potentiation. We tried to keep the neural activity in the patterns at physiologically observed levels of $1-2\%$ (Wilson, 1999).

The above changes allowed us to store around 10-times more patterns into the network than its number of neurons and the model still worked well on pattern recall process. The Hopfield model is considered in general as the appropriate model of auto-associative memories in general. Its capacity is roughly considered $0.15N$ where N is the number of neurons in the network (Wilson, 1999). Thus, we have been able to store and retrieve more than 66-times more patterns.

Experiments with the capacity showed that there is a local maximum for the potentiation parameter P at the value of 1.5, see Figure 2. This was caused by a balance between positive and negative memory tests. In general, it is good to understand that for short-term synaptic potentiation there might exist an optimal value how much the synaptic weight should be increased.

Conclusions

Our goal was to find some improvement of the existing models utilizing the dynamics of short term synaptic potentiation. We have described the artificial neural network model that uses cyclic activities as patterns that are stored and retrieved from the network. We think that these cyclic patterns are in many cases more physiological then static patterns.

We additionally extended the model with short-term potentiation which was used as vital part of the pattern recall process. We have shown that the memory capacity significantly increases in case when the ratio of active neurons is around $1 - 2\%$ which corresponds to physiologically observed values. We have done simple theoretical analysis followed by the computer simulations.

It has been shown that the short-term potentiation occurs in time range that could directly influence the pattern recall processes in biological neural networks (Tsodyks and Markram 1997, Tsodyks et al. 1998). We have shown that it is possible to find the case where these properties can rapidly improve the behavior of the artificial neural network model.

We think that there might be more cases where the synaptic dynamics of pattern recall process could improve the model performance. It might be possible that evolution managed to find more use cases for this phenomenon that still might be revealed.

We have described the artificial neural network model that uses cyclic activities as patterns that are stored and retrieved from the network. We expect that these cyclic patterns are more physiological then static patterns.

We additionally extended the model with short term potentiation which was used as vital part of the pattern recall process. We have shown that the memory capacity significantly increases in this case.

Publications of the Candidate with Co-Authors

Publications with IF

(1) Štroffek, J., Maršálek, P. and Kuriščák, E. (2007). Pattern storage in a sparsely coded neural network with cyclic activation. *BioSystems*, 89(1-3), pp. 257–263. **IF(2007)=1.646.**

(2) Štroffek, J., Kuriščák, E. and Maršálek, P. (2010). Highway toll enforcement: Real-Time classification of motor vehicles. *IEEE Vehicular Technology Magazine*, 5(4), pp. 56–65. **IF(2010)=1.184.**

(3) Štroffek, J. and Maršálek, P. (2012). Short-term potentiation effect on pattern recall in sparsely coded neural network. *Neurocomputing*, 77(1), pp. 108–113. **IF(2012)=1.634.**

(4) Kuriščák, E., Maršálek, P., Štroffek, J. and Wünsch, Z. (2012). The effect of neural noise on spike time precision in a detailed CA3 neuron model. *Computational and Mathemathical Methods in Medicine*, pp. 1–12. **IF(2012)=0.791.**

(5) Kuriscak, E. and Marsalek, P. and Stroffek, J. and Toth, P. G. (2015). Biological Context of Hebbian Learning in Artificial Neural Networks, a Review. *Neurocomputing*, 152, pp. 27–35. **IF(2015)=2.083.**

Other Publications

(1) Štroffek, J. and Kuriščák, E. (2005). Associative neural network with cyclic patterns, *Proceedings of the 5th workshop: Cognition and artificial life*, Editors: Kelemen, J., Kvasnička, V., Pospíchal, J., pp. 541–548.

(2) Maršálek, P. and Štroffek, J. (2006). Sound localization neural pathway encoding in humans *Proceedings of the 6th workshop: Cognition and artificial life*, Editors: Kelemen, J., Kvasnička, V., pp. 283–288.

(3) Štroffek J, Maršálek P and Kuriščák E (2008). Real time classification of motor vehicles in toll enforcement *Proceedings of the 8th workshop: Cognition and artificial life*, Editors: Kelemen, J., Kvasnička, V., Pstružina, K., pp. 371–323.

References

(1) R. C. Atkinson, R. M. Shiffrin (1968). Human Memory: A Proposed System and its Control Processes. *Psychology of Learning and Motivation*, 2, pp 89 – 195

(2) D. Bibitchkov, J. M. Herrmann and T. Geisel (2002). Effects of short-time plasticity on the associative memory. *Neurocomput*, pp. 329–335.

(3) D. O. Hebb (1949). The Organization of Behavior. *Wiley*, New York.

(4) J. J. Hopfield (1982). Neural networks and physical systems with emergent collective computational abilities. *PNAS*, 79, pages 2554–2558.

(5) J. J. Hopfield (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *PNAS*, 81, pages 3088–3092.

(6) W. James (1892). Psychology - The Stream of Consciousness. *Cleveland & New York*, New York

(7) N. Matsumoto, D. Ide, M. Watanabe and M. Okada (2007). Retrieval property of attractor network with synaptic depression. *J Phys Soc Japan*, 76(8), pages 084006.

(8) L. Pantic, J. J. Torres, H. J. Kappen, S. C. Gielen (2002). Associative Memory with Dynamic Synapses. *Neural Comput*, pp. 2903–2923.

(9) I. P. Pavlov (1927). Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex *Oxford University Press*, London.

(10) E. T. Rolls and A. Treves (1998). Neural Networks and Brain Function. *Oxford University Press*, New York.

(11) C. E. Shanon (1948). A Mathematical Theory of Communication *The Bell System Technical Journal*

(12) J. J. Torres, L. Pantic and H. J. Kappen (2002). Storage capacity of attractor neural networks with depressing synapses. *Phys Rev E*, 66(6).

(13) M. Tsodyks, and H. Markram (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability *PNAS*, 94(2)

(14) M. Tsodyks, K. Pawelzik, H. Markram (1998), *Neural Comput.*, 10.

(15) D. J. Willshaw, O. P. Buneman and H. C. Longuet-Higgins (1969). Non-holographic associative memory. *Nature*, 222, pages 960–962.

(16) H. R. Wilson, (1999) Spikes, Decisions, and Actions: The Dynamical Foundations of Neuroscience. *Oxford University Press*, New York