ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 97-102

Memory, Learning and Neuromediators

Alexander S. Mikhailov

Abt. Physikalische Chemie, Fritz-Haber-Institut der MPG,
Faradayweg 4-6, 14195 Berlin, Germany, and
N.N.Semenov Institute for Chemical Physics, Russian Academy
of Sciences, ul. Kosygina 4, 117334 Moscow, Russia

Abstract. We consider a model of a neural network where individual cells interact only by releasing and receiving molecules of a neuromediator. We show that such a system can realize the function of associative memory. The mechanisms of learning involving neural plasticity are discussed.

1. Introduction

The neural systems of the brain differ significantly from the idealized networks of formal neurons. Real neurons are surrounded by the medium that contains various types of chemical agents secreted by the cells. The chemical composition of this medium and the spatial distribution of the mediators in it are permanently changing. These changes may influence the neural activity including the secretion process itself. Hence, besides of the direct connections through synapses, the neurons can also interact by generating and receiving different chemical signals.

If this view of a neural network is accepted, the central question is what might be the role played by chemical mediators in information processing. It was conjectured some time ago by Shepherd [1] that exchange by chemical signals can effectively lead to the same kind of information interactions that is realized by direct synaptic connections. Recently we constructed and examined several mathematical models [2-6] of neural networks where direct synaptic connections were absent and the interactions between neurons involved mediators (or messengers).

The first of these models [2,5-6] used addressable messengers which could act only on the neurons with an address that coincided with the one carried by the messenger. It was shown that, under a proper choice of the communication rules, this system can emulate any activity of a formal neural network. In the second model [4-6] each mediator was associated with a particular stored pattern and the system had the property of associative memory. In both of them diffusion of the mediators was assumed to be so fast that their concentrations were maintained uniform in the medium.

In the present paper (see also [3,6]) we investigate another model of a neural network. Its principal difference is that it uses a single kind of a mediator while retaining the potential to keep and to recall a large number of patterns. We formulate further a possible learning mechanism based on the processes of neural plasticity. The theoretical analysis of the model is supplemented by numerical simulations.

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 97-102

2. The model

We neglect direct connections between neurons which might be realized by synapses. Instead it is assumed that a neuron can only interact with the others through the intracellular medium. At its terminals located on the axon, the neuron releases into this medium the molecules of a mediator. This mediator diffuses in the intracellular space and acts on the dendritic receptors of other neurons. We assume that the dendritic tree of any neuron is strongly branched and has a very large number of receptors. Then, in the simplest approximation, one can introduce the local density of receptors of a given neuron which is described by a continuous function of spatial coordinates. The axons are also assumed to be strongly branched. For every neuron, we define its local continuous density $A(\mathbf{r})$ of axon terminals where the mediator is released. Thus, a neuron is pictured in our model by the clouds of its dendritic receptors and axonic terminals.

In the brain there are dozens of chemical substances which act as neuromediators. In our model, however, all interactions between the neurons are realized only by a single chemical mediator. This simplification can be partly justified when we note that the neuromediators found in the brain have very different life-times. Therefore they are involved in the processes which are characterized by different time scales and are thus sufficiently independent.We denote the local concentration of the mediator by $Q(\mathbf{r},t)$.

We suppose that the neurons (enumerated by index i) possess two types of receptors, activatory and inhibitory. Hence the receptive field of a neuron is defined by two distributions, i.e. by the local density $R_{1,i}(\mathbf{r})$ of its activatory receptors and by the density $R_{2,i}(\mathbf{r})$ of its inhibitory receptors.

In the presence of the mediator, an activatory receptor increases the electric potential U_i of the ith neuron by a contribution that is proportional to the mediator concentration Q at the location of this receptor. Under the same conditions, an inhibitory receptor decreases the potential U_i by the amount proportional to Q. We neglect the delays due to the finite speed of propagation of elecrical signals inside neurons and assume that at time t the electric potential of a neuron integrates over the instantaneous contributions from all the receptors, i.e.

$$U_i(t) = \int \alpha[R_{1,i}(\mathbf{r}) - R_{2,i}(\mathbf{r})]Q(\mathbf{r},t)d\mathbf{r} \qquad (1)$$

Here α is a positive coefficient.

We describe the dynamics of the neural activity in the simplest form, assuming that the mean frequency ν_i of spikes generated by the ith neuron is a step-like function $\nu_i = f(U_i)$ of the potential U_i. The monotonous function $f(U)$ vanishes at $U = 0$ and sharply increases when a threshold U_c is exceeded. At large values of U this function reaches saturation. Below ν_i is called the activity of the ith neuron.

To complete the model, we provide an equation describing evolution of the mediator distribution in the medium,

$$\partial Q / \partial t = D\nabla^2 Q - \gamma Q + \beta \sum_i f(U_i)A_i(\mathbf{r}) + I(\mathbf{r},t). \qquad (2)$$

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 97-102

The first term in the right side of this equation takes into account diffusion of the mediator, the second describes decay of the mediator and its absorbtion by the cells (neurons and glia). The third term decribes release of the mediator by the neurons at their axon terminals. It is assumed here that the amount of the mediator reeleased by each neuron is proportional to its instantaneous activity $\nu_i = f(U_i)$. The last term takes into account possible sources which might locally feed the mediator into the system.

When distributions $R_{1,i}(\mathbf{r})$, $R_{2,i}(\mathbf{r})$ and $A_i(\mathbf{r})$ are fixed, equations (1) and (2) govern time evolution of the mediator distribution $Q(\mathbf{r},t)$.

3. Associative memory

With a proper choice of the spatial distributions of receptors and terminals of the neurons comprising the system, it can keep in its memory a set of patterns. Such patterns represent different spatial distributions of the mediator. Each prototype pattern is uniquely associated with a particular neuron. If some initial disribution of the mediator is applied, the system evolves into a stationary state with the mediator distribution of the nearest prototype pattern. When this occurs, the neuron associated with this prototype pattern becomes activated while all other neurons are silent (i.e. their activities are vanishingly weak).

It was shown in [3] that, in order to keep in the memory a set of N prototype patterns $Q_i(\mathbf{r})$, $i = 1, ..., N$, the spatial distributions of dendritic receptors and axonic terminals of ith neuron can be chosen as

$$R_i(\mathbf{r}) \propto \delta Q_i(\mathbf{r}), \qquad A_i(\mathbf{r}) \propto Q_i(\mathbf{r}). \tag{3}$$

Here $R_i(\mathbf{r}) = R_{1,i}(\mathbf{r}) - R_{2,i}(\mathbf{r})$ is the local difference in the densities of the activatory and inhibitory receptors of the ith neuron and $\delta Q_i(\mathbf{r})$ is the local variance for the ith prototype pattern, defined as $\delta Q_i(\mathbf{r}) = Q_i(\mathbf{r}) - \langle Q_i \rangle$ where $\langle Q_i \rangle$ is the spatial average for this pattern.

The underlying mechanism of this behaviour is simple: If the receptive field $R_i(\mathbf{r})$ of a certain neuron is closer approaching the presented pattern $Q_0(\mathbf{r})$, its initial activation is higher and, consequently, it begins to release the mediator at a higher rate. But, according to (3), the spatial distribution of its axonic terminals repeats, up to a constant component, the pattern of the receptive field. Therefore, the applied pattern evolves towards the prototype pattern kept by this neuron, until this pattern is fully established in the medium. When several neurons become initially activated, the competition between them starts. In its course, the neuron with the best fit between its receptive field and the initially presented pattern suppresses the activity of all other neurons and the mediator distribution, generated by this neuron, appears in the system.

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 97-102

4. Learning

The neurophysiological observations provide evidence that learning may be related to neural plasticity, i.e. to the processes which involve growth (sprouting) of neurons, establishing new synaptic connections and modifying the strengthes of already existing ones. The last two effects, however, cannot play a role in the considered model because the synaptic connections between the neurons are excluded here.

It is known that sprouting of neurons is controlled by trophic factors which represent certain chemical substances. The best studied trophic factor is a protein, the nerve growth factor (NGF). This protein stimulates directed sprouting of axons and dendrites by control of assembling the microtubules. The axons and dendrites adjust their directions of growth to the concentration gradient of NGF. According to a popular hypothesis, the peripheral target cells produce NGF which directs sprouting of axons to their targets. Hence, sprouting represents a chemotaxis-like process (with the difference that not the entire cells but only their parts move in the direction of a chemical gradient).

Although there is no direct evidence, it seems plausible that motion of receptors and of mediator terminals on the membranes of neural cells may also be controlled by chemical gradients. This would mean that such receptors and terminals are involved in chemotaxis (confined however to the cellular membrane of a given neuron).

As already noted, our simple model is not intended to give a complete description of a real physiological system. Since we want only to demonstrate the possibility of such mechanism of learning, we assume for simplicity that the mediator itself is a substance which controls all plastic changes in the neurons and their sprouting.

In our approach any neuron is characterized by a set of three fields giving the local densities of dendritic receptors and axonic terminals. These individual receptors and terminals can be viewed as some particles which are distributed over the medium. Such particles can move, either together with a sprouting axon or dendrite to which they belong, or over the membrane of a resting neuron. Abstracting from the details of such motion, we describe it phenomenologically as chemotaxis of these particles controlled by the mediator gradient.

We assume that axon terminals tend to drift in the direction of the decrease of the mediator concentration. The drift occurs only when a given neuron is active. Dendritic activatory receptors drift in the direction of the increase of the mediator concentration, if a given neuron is active, and in the opposite direction if its is currently passive. The drift is superimposed on random diffusional wandering of these particles through the medium. As for inhibitory dendritic receptors, we assume that they are uniformly distributed in the medium and their distribution is not changed in the process of learning.

The chemotaxis-like motion of activatory dendritic receptors and axonic terminals is described by equations

$$\partial R_{1,i} / \partial t = D_i \nabla^2 R_{1,i} - \mathrm{div}(p_1(v_i) R_{1,i} \mathrm{grad} Q) \quad , \tag{4}$$

$$\partial A_i / \partial t = D_a \nabla^2 A_i - \mathrm{div}(p_a(v_i) A_i \mathrm{grad} Q). \tag{5}$$

ESANN'1994 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 20-21-22 April 1994, D-Facto public., ISBN 2-9600049-1-4, pp. 97-102

Here D_1 and D_a are the coefficients that characterize the rates of undirected diffusion-like spreading of dendritic and axonic clouds in absence of the mediator gradient. Coefficients p_1 and p_a characterize plastic sensitivity of dendrites and axons to mediator gradients. They depend on the current activity $v_i = f(U_i)$ of a given neuron,

$$p_1(v_i) = k_1(v_i - \theta) \quad , \qquad p_a(v_i) = k_a v_i \quad . \tag{6}$$

Thus, the activatory receptors drift in the direction of the mediator increase when activity of the neuron exceeds threshold θ. When $v_i < \theta$ the direction of their drift is reversed. The axonic terminals always drift only in the direction of the gradient.

The learning procedure is the following. Suppose that we have a set of different distribution patterns of the mediator and want to teach the ith neuron to recognize only a particular pattern characterized by the mediator distribution $Q_i(\mathbf{r})$. Then we create this mediator distribution in the medium and maintain it fixed for some time. During this time we keep the ith neuron in the active state (with $v_i > \theta$) and all other neurons in the passive states. Under these conditions, the activatory receptors of the ith neuron drift towards the areas where the mediator concentration is maximal, while the activatory receptors of all other neurons tend to leave such areas. The axonic terminals of the ith neuron go into the regions with the higher mediator concentrations whereas the terminals of other neurons do not perform any directed motion.

The training cycle consists of application in turn of all prototype patterns, each kept for a short time interval to produce changes in the dendritic and axonic fields. It is expected that after repetition of many training cyles the system acquires the ability to recognize and reproduce stored patterns.

5. Numeric simulations

To test the efficiency of the learning procedure, numerical simulations were performed. In the first of them we trained the system, consisting of two neurons, to recognize patterns which represented letters *L* and *U*. The details of this simulation are given in [3]. After the learning procedure was finished, we used this system to recognize the prototype patterns *U* or *L* in the presented pictures. The applied pattern, which was used to create the initial mediator distribution, represented a strongly distorted letter *U* . We were able to see how the missing elements appeared and the superficial elements faded out with time. The final mediator distribution reconstructed the full prototype pattern.

The second computer experiment (see [3]) differed in the number of patterns that were learned by the system. The set of the prototype patterns consisted now of four digitazed photographs of different human faces.The learning procedure which was employed in this simulation has been slightly modified.

As noted above, the equations governing evolution of axonic and dendritic distributions in the process of learning include diffusion-like terms. These terms describe undericted sprouting of neurons and random wandering of the receptors and terminals over the cellular membranes. The rates of such random spreading are

characterized by two diffusion constants D_1 and D_a. The presence of random motion is essential for efficient learning.

Suppose that initially the axonic and dendritic distributions are localized inside a certain spatial region. Then, in order to produce distributions that correspond to the prototype patterns occupying larger space regions, the initial clouds of receptors and axonic terminals must spread over the medium. However, it would be very often found that such spreading could not result from the drift along the mediator gradients. Indeed, to reach a distant region where their local maximum should be established, the receptors and the terminals must first pass through the regions where the local mediator gradients might be looking in the opposite direction. Arriving into such areas can occur only by chance, i.e. as a result of random sprouting.

At the initial stage of learning the diffusion constants should be sufficiently large. The intensive random initial sprouting allows the dendritic and axonic clouds to spread over the medium and to reach the areas where local maxima of receptive and generating fields should be later established. On the other hand, random spreading smears the details of patterns. If the diffusion constants are kept large, the evolving dendritic and axonic fields would catch the rough features of the learned patterns but would not be able to resolve the fine structure of these images.

To satisfy both conditions, a procedure similar to the technique of simulated annealing has been used in the computer experiment. At the initial stage of learning, higher values of the diffusion constants D_1 and D_a were chosen. They were then gradually diminished in subsequent training cycles so that the finer details of the digitazed photographs have been learned.

Thus we have shown that a system of neural cells interacting only by release and reception of a single mediator substance is already able to perform complicated tasks of information processing and that the mechanism of learning based on neural plasticity can operate even in absence of direct synaptic connections between the cells.

References

1. G.M. Shepherd, *Neurobiology* (Oxford Univ. Press 1983)
2. A.S. Mikhailov, I.V. Mit'kov and N.A. Sveshnikov, BioSystems 23, 291-295 (1990)
3. E.M. Izhikevich, A.S. Mikhailov and N.A. Sveshnikov, BioSystems 25, 219-229 (1991)
4. A.S. Mikhailov, *Foundations of Synergetics I. Distributed Active Systems* (Springer, 1990)
5. A.S. Mikhailov, in *Rhythms in Physiological Systems*, eds. H. Haken and H.P. Koepchen (Springer, 1991) pp. 339-350
6. A.S. Mikhailov, in *Interdisciplinary Approaches to Nonlinear Complex Systems*, eds. H. Haken and A.S. Mikhailov (Springer, 1993) pp. 89-108