

On a Simple Stochastic Neuron – Like Unit

W. Banzhaf

Institut für Theoretische Physik und Synergetik der Universität, Pfaffenwaldring 57/IV,
D-7000 Stuttgart 80, Federal Republic of Germany

Abstract. We consider a simple stochastic unit realized by digital ANDs and ORs. The function of the unit is inspired by nerve cells found in brains of higher organisms. Information is carried by trains of pulses in time. Therefore, time is introduced into the model as an essential variable. Principles of stochastic computing are used to appropriately model weighted summation of inputs. The model unit allows to process analog information as is provided by observables of real environment. The statistical properties of the model are examined. The traditional saturation non-linearity of neurons emerges as a natural consequence of signal gating by “synapses”. Different schemes of synaptic modification are indicated.

1 Introduction

The starting point of our considerations is the fact that nervous systems use frequency coding of signals which was discovered some 50 years ago (Adrian 1932; Stein 1967; Eccles 1973; Bullock et al. 1977). The concept of frequency coding, however, is only applicable to nervous tissue “cum grano salis” since a great diversity of neurons exists generating a whole spectrum of signal codes. Neurophysiologists even claim (Bullock et al. 1977) that each neuron has its own individuality setting it apart from all other neurons.

Any code which uses the number of pulses in certain time intervals as information carrier is a “non-valued” computing system. In general, such systems possess much more fault tolerance than valued computing systems as e.g. a number system differentiating their values at different positions.

Despite the fact that neurons and nets of neurons were modeled as all-or-none units over many decades (McCulloch and Pitts 1943; Steinbuch 1960; Hopfield 1982) they are not at all binary valued elements. But as

computer scientists learned during the development of artificial information processing systems, a binary coding of information is very fault tolerant also. The true discovery of digital age, however, is not the binarity of computer’s elements, but their non-linearity. Indeed, this was “invented” some millions of years ago by nature during the evolutionary process leading to neurons which are equipped with different non-linearities, the most important of which is called the saturation non-linearity. Since nervous systems have to process analog quantities (INPUT = analog observables measured by sensors, OUTPUT = analog actions executed by actors), it is not astonishing that an electro-chemical coding was developed during evolution which unites the advantages of binary and analog coding schemes. Pulses are a very secure form of signal transfer which is a far reaching benefit in systems with long signal ways such as nervous systems. Moreover, the huge amount of computation is to be executed in elements which are to a certain degree not reliable.

Finally, one may hypothesize that a little ingredient of stochastics in nervous systems as it is provided by some noise from unreliable elements and the indeterminism originating from probabilistic processing is by no means an evolutionary accident but a precise reflection of certain environmental conditions which otherwise would have been very difficult to catch.

Here we want to present first results concerning the simulation of a neuron-like unit. It was modeled using logical gates which are nowadays available in large numbers on single VLSI chips. The unit is able to “integrate” the incoming signals, a function which is at least in part analogous to that of real neurons. The model is intended to proliferate an instrument for examination of questions related to the collective behavior of certain networks.

We do not claim to have a realistic model of biological nerve cells but to use similar stochastic

computation methods as real neurons do (Von Neumann 1956; Ribeiro 1967; Peretto and Niez 1986; Gaines 1987).

2 The Coding of Signals

To motivate our proceeding below different arguments can be given:

i) The Parallel Processing Argument

In a system where information is carried by pulse sequences stochastically distributed over time parallel processing is possible. In fact, a continuous enhancement of accuracy is reached with every addition of computational elements. The more neurons code the same incoming signal and afterwards process it, the more exact the result of a computation will be. In the language of physicists the argument reads: One can substitute the time average (of a single measurement or computation) by an ensemble average (of multiple measurements or computations), if the underlying process is ergodic. Let us suppose for the moment that the processes in nervous systems are sufficiently independent that the hypothesis of ergodicity is valid. If we turn the argument around: The same measurement accuracy will require a lower measurement duration if we use more elements.

ii) The Energy Consumption Argument

Bearing in mind the very efficiency of nature one is led to the assumption that an effective operation will mean for the neurons *not* to use the extremal states too often. Thus, a tendency to avoid the "0" and "1" states of firing at lowest and highest rates possible is claimed. It is further supposed that activities lie in general at the lower end of the scale. The reason is that every activity is costly in terms of energy and that, additionally, release of transmitter substance in synapses consumes ATP. This argument is valid for excitatory *and* inhibitory synapses.

iii) The Compatibility Argument

Pulses are transmitted from sensory units to other processing units presumably working according to the same principles. To allow further processing of initially fuzzy input information without compatibility problems all units should use nearly the same coding. Especially, layered structures would need the same analog-processing capabilities in deeper layers.

It is fascinating to follow these arguments in constructing a technical system applying similar computational methods under different circumstances. In fact, we shall follow the above lines of thought.

Therefore, let us introduce stochastic pulse trains which will code analog quantities s in probabilities that our neuron-like unit will fire or not fire in a certain time interval T :

$s \rightarrow P(s)$: firing probability in T .

The detailed dependence could be a linear, logistic or even a logarithmic function of signals s depending on circumstances. The only condition to be satisfied is that $P(s)$ is an ever increasing function of s .

The first place where such a coding may occur is the sensory input to a net of these neuron-like units. The detailed mechanism converting mechanical, optical or other stimuli into electro-chemical pulses will determine the specific shape of $P(s)$. "Hidden" units not receiving information from the outer world may be interpreted as sensory units for electro-chemical potentials such as those generated near the axon hillock of an ordinary neuron.

Every functional dependence will result in a different shape of the pulse partition over time. Considered over long enough times a sort of binomial distribution may be generated:

$$B(k) = \binom{N}{k} P(s)^k (1 - P(s))^{N-k}, \quad (1)$$

where time is quantized and $B(k)$ represents the probability that in a sample of N elementary time steps k pulses are realized. The minimum number of pulses in T is 0 whereas the maximum number is given by the dead-time of the units and will be called N . Therefore, the real numbers $s \in [0, 1]$ are mapped onto the numbers of pulses in time intervals T . Equivalently, the time intervals between pulses will have different distributions. If time is not quantized pulses are distributed according to a Poisson distribution.

Four possible codes are reasonable:

- i)* a pure frequency code
- ii)* a stochastic frequency code
- iii)* a pulse-distance code
- iv)* a stochastic pulse-distance code.

For specific problems, all of these codes may have advantages, and it is a matter of intuition which one to choose.

3 Application of Digital Technique

We now want to make contact with digital technique. Today, this is a widespread and handy technology realized physically by different methods. We shall see an interesting relation between both of our subjects. The point is that we can use digital pulses as working material in artificial neuronal circuitry.

First of all, however, we have to make the approximation that dispersion in our system is not present

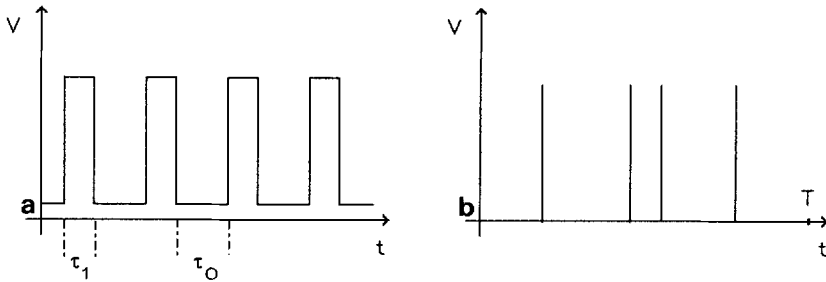


Fig. 1a and b. Illustration of digital signals used. V could mean voltage differences as well as other kinds of physical quantities possibly used.
a General pulses without dispersion.
b The approximation used: $\tau_1 \ll \tau_0 \ll T$

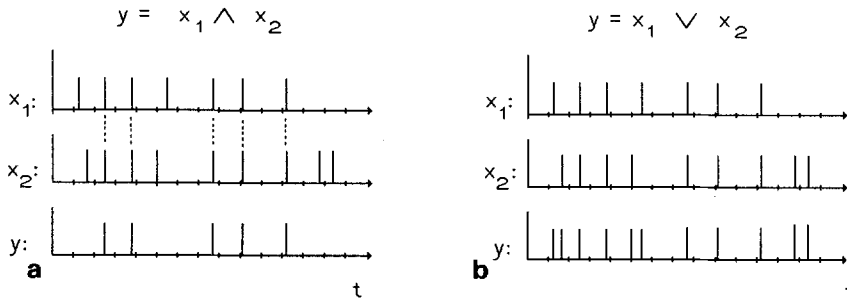


Fig. 2. a AND as coincidence measurement, **b** OR as unification of pulses on channels x_1, x_2

(Fig. 1a). This is a remarkable difference to conventional computers which are static machines suppressing all their dynamics by tact cycles. In contrast, here we have a truly dynamical concept of pulse processing. We use pulses with $\tau_1 \ll \tau_0$ and consider fixed time intervals of length T , $T \gg \tau_1, \tau_0$ (Fig. 1b).

Furthermore, digital technology makes available logic gates of different sorts. For our purposes the AND and OR gates are adequate. If we use a stochastic coding, AND can be seen as the measurement of coincidence between signals and amounts to a multiplication of analog quantities $s_1 \cdot s_2$ corresponding to pulses x_1, x_2 , whereas OR means unification of signals or roughly addition $s_1 + s_2$ (see Fig. 2a, b).

The principles of stochastic computation are well established so we only refer to the literature on the subject (Gaines 1969; Massen 1977). In contrast to the traditional approach to stochastic computing, our attention will not be directed towards an accurate realization of algebraic operations but towards usage of gates to realize primitive neuron-like functions. A proper arrangement of gates is shown for a specific example in Fig. 3.

The circuit is designed so as to compute roughly

$$s^{out} \approx \sum_{i=1}^5 w_i \cdot s_i^{in} \quad (2)$$

expressed in the corresponding analog quantities.

It is now evident that stochastic events, i.e., pulses partitioned at random on channels $x_1 \dots x_5$ are necessary to approximate the summation function of the OR gate. Thus, randomness is utilized and we do not

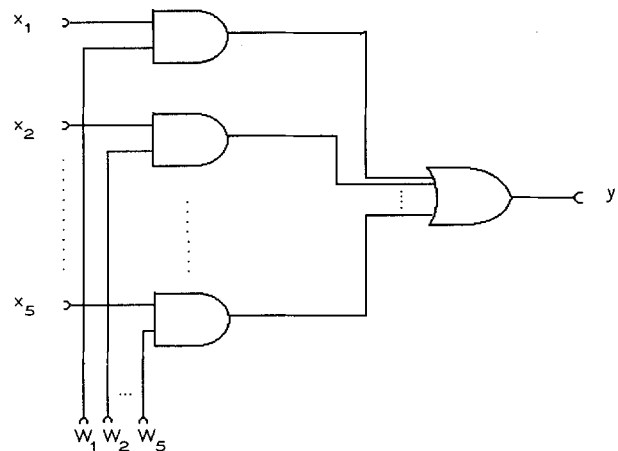


Fig. 3. Arrangement of digital gates that operates in a nearly neuron-like fashion. Input channels are $x_1 \dots x_5$, (synaptic) weights are $W_1 \dots W_5$. Output is y

struggle against noise. We suppose that a code like iv given above will give the best result with respect to an approximation to the sum in s^{out} .

On principal, channels x and W are symmetrical. We shall, however, agree upon the following discrimination: The W -channels transferring stochastically distributed pulses represent analog quantities called synaptic weights w . These will have a dynamics obeying another time scale which is much slower than that of the signals s behind x -channels: $\tau_w \gg \tau_s$. Figure 4 gives an illustration of the signal processing in a single unit. Without manipulating weights, a non-linearity in the input-output-function of the unit emerges (cf.

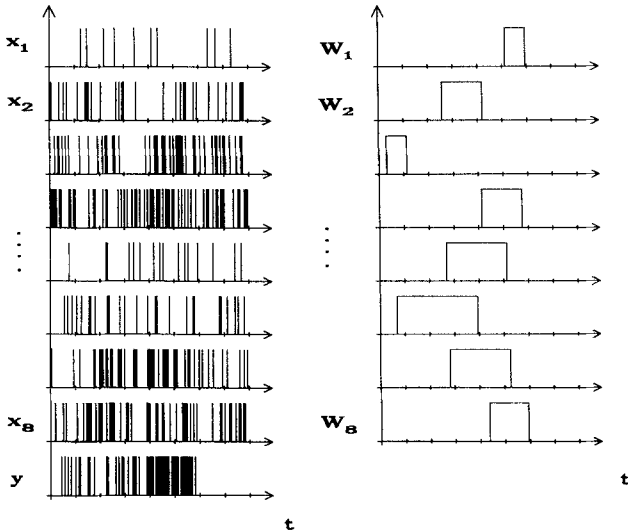


Fig. 4. Processing of pulses during a time interval T . Input channels x are gated by “synaptic” channels W . The result is shown on channel y

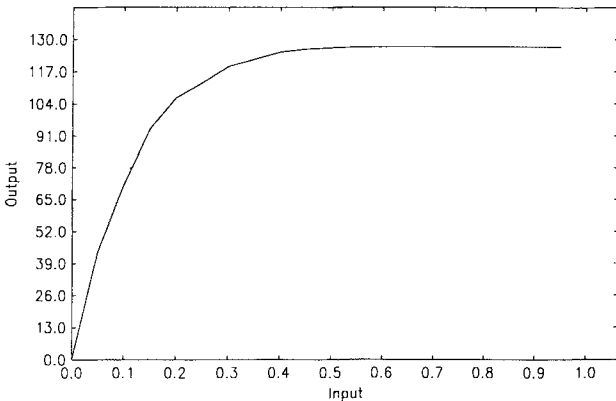


Fig. 5. Input-output function of a single unit with fixed weights. Ordinate axis and abscissa are “firing” probabilities in input output channels respectively

Fig. 5). Clearly, this is the saturation effect of the output channel after summation due to its fixed dead-time!

Let us make some comments on stochastic weights W_i :

a) We have chosen a slower time scale for W -channels under the aspect of adaptivity of the system. It is a well established fact that two different time scales in a dynamical system can cause adaptive properties.

b) The weights w_i act as gating probabilities for pulses x_i and are therefore taken from $w_i \in \mathbb{R}$, $w_i \in [0, 1]$. The gating may be realized by frequency manipulation or by tuning of puls-broadness. Altogether, the result is a certain probability of gating the pulses on x -channels.

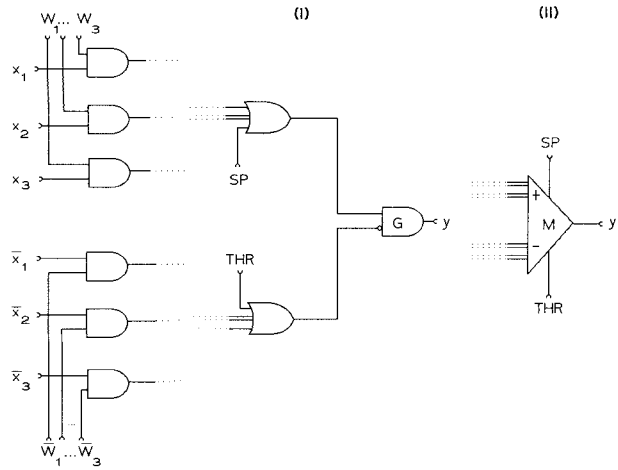


Fig. 6. Two ways of combining excitatory and inhibitory channels of the system. (I) with a special gate G . (II) with a majority decision element M

c) The above view allows one important possibility: W_i could operate as a filter over time which allows passage of certain correlated pulse trains whereas it prevents others. One can think of circumstances, when small weight probabilities let pass all the signals in a train, and on the other hand, when no signal can pass although weights are considerably high. This is a direct consequence of treating time as an essential variable in our model! In that respect, a description of neuronal net dynamics just using numbers between – let’s say – 0 and 1 is misleading.

d) A special mode of operation is working close to an instability point. In that case a minimal change in the weight probabilities results in a radical change of the overall behavior of the unit.

e) Comparing the proposed unit to real neurons we can relate the weights w to probabilities of releasing vesicles of certain neurotransmitters into the synaptic cleft. Control is executed by variation of the respiration time a synapse needs to be able to react again.

Before coming to learning, i.e., the dynamics of the weights, we want to present a more realistic system having still more similarity to neurons. To this end we use inhibitory channels which are arranged just as the excitatory channels of Fig. 3. We combine inhibitory and excitatory channels and finally add some spontaneous activity on both of these. The inhibitory spontaneous activity will be called the thresholding. Spontaneous activity is fed into the net without any gating or, in other words, with a fixed “1” as gating probability. Figure 6 shows two alternative ways of combination (I) and (II)). We end up with a very familiar I/O

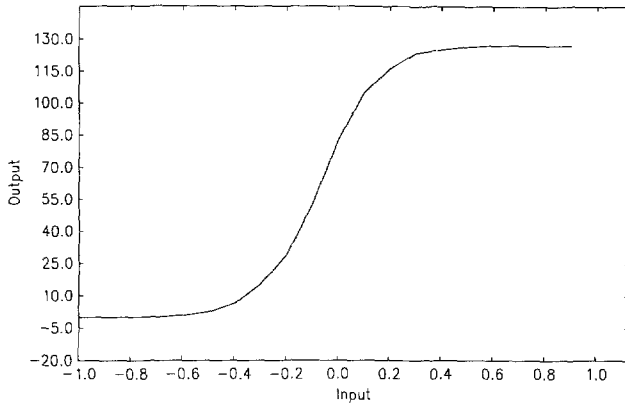


Fig. 7. Input-output relation of unit of Fig. 6. Negative inputs mean inputs on inhibitory channels

relation known as the saturation non-linearity of a neuron (cf. Fig. 7).

Now we have realized some of the four loci which are found in neurons (Lewis 1983):

- 1) Synaptic loci: inputs from other neurons, separated in inhibitory and excitatory ones.
- 2) Pacemaker loci: spontaneous signal generation in the neuron itself.
- 3) Local response loci: Signal integration inside the neuron.
- 4) Spike initiator loci: Transformation of analog signals into spike trains.

In case (I) of Fig. 6 the first and second, in case (II) all loci are present. The saturation non-linearity of nerve cells is seen to arise quite naturally. It could be conjectured that a low level of activity in a network of units of this type will be favourable. If one wishes to include higher order correlations this may be realized just using AND gates with more than two input lines.

4 Statistical Behavior of the Unit

We consider signals in the interval $s \in [-15, +15]$ using, for the sake of simplicity, the linear dependence

$$p(s) = \alpha \cdot s + \beta.$$

Two alternative ways of generating pulses from signals arise:

- (a) direct generation
- (b) Gaussian generation.

Technically, the former consists of drawing a random number $z \in [0, 1]$ at every elementary time step $i \cdot \tau$, $i = 1, \dots, N$, $\tau = T/N$ and comparing $p(s)$ with z . A pulse is then released every time when a number z smaller than $p(s)$ was drawn.

In contrast, the latter possibility amounts to summing up random numbers z_i over time (indexed now for different time steps i). A pulse is released only if

$$\sum_i z_i \cdot p(s) \geq 1. \tag{3}$$

Then, the whole process begins anew.

We call this sort of pulse generation Gaussian, because a well known random number generator (generating white noise) implements just this method (Hemmerle 1967). Figure 8 shows three different samples generated by either method.

If one plots $\langle \Delta t \rangle$ over $p(s)$, a geometric distribution of pulse distances becomes visible. This is an interesting result which needs further explanation.

In the following, we abbreviate the signal probability by p and its complement by $q = 1 - p$. In a random sample with the maximum of N possible elementary events the number of realized pulses will be distributed according to (1). As a consequence, the average number of pulses $\langle k \rangle$ will be

$$\langle k \rangle = N \cdot p \tag{4}$$

3 DIFFERENT SIGNALS CODED

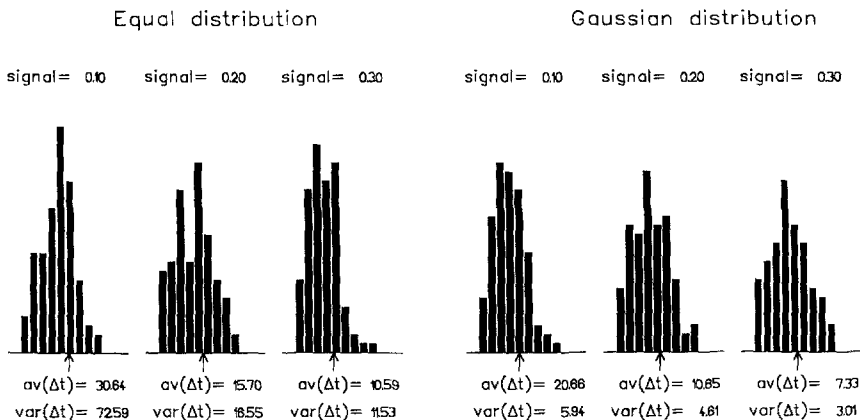


Fig. 8. Three different signals translated into pulse trains of 1000 pulses. The relative distribution of pulse distances Δt is shown for both methods of pulse generation. Average and variance of distances is sketched also

Table 1. Comparison of the summation operation (O) of the unit (cf. Fig. 3) and an algebraic summation (Σ) of the corresponding quantities. Shown is the deviation $\text{Dev} = \left| \frac{O - \Sigma}{\Sigma} \right|$ for $\langle s \rangle = 0.001$ and $\langle w \rangle = 0.5 + N(0, 0.05)$

Average ov. time steps	Number of inputs N_I							
	5	10	20	50	100	200	500	1000
10	1	1	1	1	5.52	1	0.16	0.39
50	1	1	1	1	1.17	0.25	0.01	0.07
100	1	2	1	0.54	0.52	0.06	0.03	0.07
500	1	0.4	1	0.09	0.35	0.04	0.12	0.12
1000	0.43	0.2	0.12	0.04	0.11	0.07	0.14	0.2

and the variance σ reads

$$\sigma = \sqrt{\langle k^2 \rangle} = \sqrt{Npq}. \quad (5)$$

Both of these quantities are dependent on the size of the sample. On the other hand, the time between two pulses obeys a "negative" binomial distribution (Feller 1976). This is the probability distribution for waiting times between the n -th and $(n+r)$ -th pulse which is sometimes also called Pascal distribution:

$$f(l; r, p) = \binom{-r}{l} p^r (-q)^l = \binom{r+l-1}{l} p^{r-1} q^l \quad (6)$$

times p . Since we are interested in the waiting times between a pulse and its subsequent one we have to set $r=1$. The result is the geometric distribution

$$G(l) = p \cdot q^l. \quad (7)$$

The average waiting time reads

$$\langle l \rangle = qp(1 + 2q + 3q^2 + \dots) = qp(1 - q)^{-2} = \frac{1}{p} - 1 \quad (8)$$

which is independent of N . Similarly,

$$\sigma = \sqrt{\langle l^2 \rangle} = \frac{\sqrt{q}}{p} = \sqrt{\frac{1}{p^2} - \frac{1}{p}}. \quad (9)$$

The geometrical distribution – which changes into the exponential distribution in the continuous time limit – is a consequence of the assumption that different random events on channels are statistically independent. As a result, the subsequent events do not influence each other.

Since s_i^{in} , w_i are probabilities, s^{out} is also a probability corresponding to the pulse partition on the output channel y . Thus,

$$s^{\text{out}} \leq 1 \quad \text{or} \quad \sum_{i=1}^{N_I} s_i^{\text{in}} \cdot w_i \leq 1.$$

Under the simplified assumption of

$$s_i \rightarrow \langle s \rangle = \frac{1}{N_I} \sum s_i \quad \text{and} \quad w_i \rightarrow \langle w \rangle$$

saturation sets in when

$$\sum_i s_i \cdot w_i \approx \sum \langle s \rangle \cdot \langle w \rangle = N_I \langle s \rangle \langle w \rangle \geq 1.$$

For $\langle w \rangle \approx 0.5$ and signals $\langle s \rangle \sim 0.1$ this leads to $N_I \geq 20$.

Vice versa, for $\langle w \rangle \approx 0.5$ and $N_I = 1000$ only an average activity of $\langle s \rangle = 0.002$ is needed to saturate the unit. Under these circumstances the whole range of different behavioral patterns is accessible. Table 1 compares a real summation of signals and the operation of the unit for different numbers of inputs.

5 Learning

An interesting aspect of the model can be seen by considering different schemes of learning. Here, we want to restrict ourselves to single neuron learning. In addition, we want to consider only non-Hebbian learning rules. The adaption of a neuron to its input data can be illustrated very clearly (Omohundro 1987): In the input space

$$\sum_i s_i \cdot w_i = \Theta \quad (10)$$

defines a hyperplane which can be shifted by changing w and the threshold Θ . If a signal lies inside the hyperplane, the neuron fires with the corresponding output O . Otherwise it is silent (no spontaneous firing for the moment). Learning means displacement of the hyperplane towards the best sensibility region to fulfill a given task.

The simplest learning process would be comparable to the process of growth: A synapse which is used often (on the average in time) becomes more "transparent"

to signals. This is equivalent to the principle that stress leads to a gradual strengthening of the living tissue involved (Thomson 1917). Note that Hebb's learning rule (Hebb 1949) uses a more sophisticated version of growth: In addition to the incoming signal on a synapse, the overall activity of the postsynaptic cell is taken into account.

The simple form of learning may be implemented using the following algorithm: The synapse i just adapts to the present signal $s_i(t)$ as

$$w_i(t) = \sigma[\gamma \cdot w_i(t-1) + (1-\gamma) \cdot s_i(t)]. \quad (11)$$

σ is a saturation non-linearity, γ may decrease slowly with time $\gamma = \gamma(t)$.

This process changes the probability of gates to let pass signals. Under these conditions, a synapse adapts slowly to the average signal strength.

Very interesting effects, however, are possible by adjusting the on-off states of synaptic gates *without* changing the overall probability of the synapse to let pass pulses. Technically one can speak of a different partition of the on-area of a synapse in T . Figure 9a indicates the effect of partitioning the gating time in different ways.

Promising are as well variable respiration times of synapses which could be used to gate specially correlated pulse trains (Hasenburg 1987). Let us define a dead-time d of a synapse in the following way: After a pulse has arrived at a synapse S at time t_0 and has then passed the synaptic gate at $t = t_0 + \tau_1$, the gate is closed during the dead-time d . If a pulse is arriving during that time $t_0 + \tau_1 \leq t_1 \leq t_0 + \tau_1 + d$, no postsynaptic pulse leaves the gate. The sensibility of the synapse, however, is not turned off. Thus any pulse arriving at S will have an effect on it. Namely, it resets the synapse and starts the respiration cycle again leading to a longer respiration time $t_2 = t_1 + d$. The situation is sketched in Fig. 9b. As a consequence, a pulse coming too early is not transferred to the postsynaptic unit and, additionally, it lengthens the closing time of the synaptic gate.

Thus, a sharp transition of the effective synaptic weight defined as

$$w_{eff} = s_{post}/s_{pre}$$

may be expected when the presynaptic signal is changed slowly. The simulation is shown in Fig. 10 for four different values of the dead-time.

A sort of habituation emerges if synapses tend to fix their dead-times such that effective weights near 1/2 arise. Depending on input and network topology, this instability point may be reached in a slow process. Unfamiliar signals quickly change the effective weights and the network as a whole makes a transition into a totally new state.

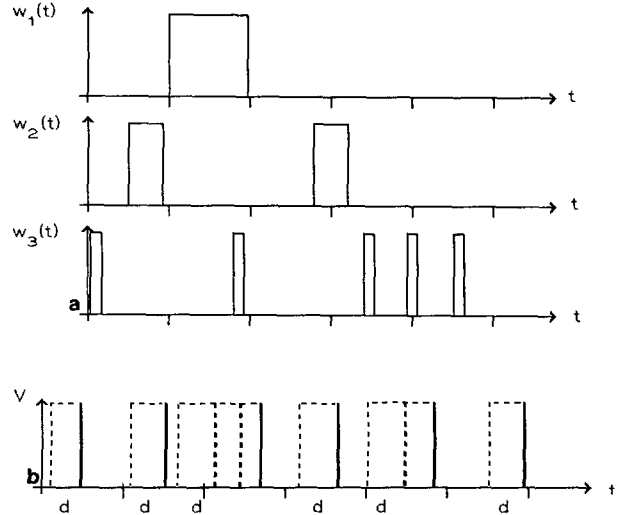


Fig. 9. a Three different partitionings of the same weight. **b** Effect of dead-time d on pulses arriving at a synapse: during the time marked by dashed lines no pulse is gated. The dashed pulses only lengthen the dead-time

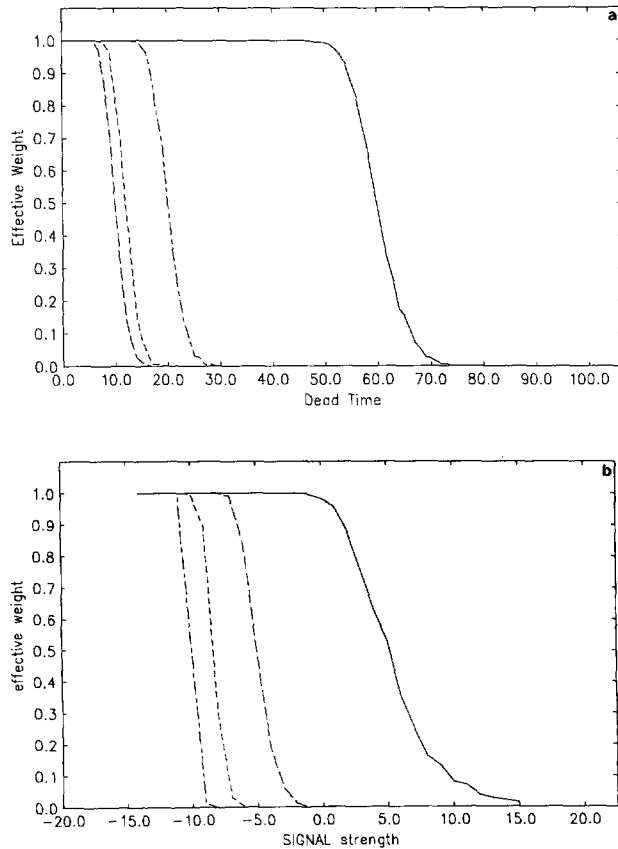


Fig. 10a and b. The effective weight w_{eff} of a synapse. **a** Changing dead-times for different signals: —: -10, ---: 0, -.-: 10, ···: 15. **b** Varying signals for different dead-times (in elementary time steps): —: 15, -.-: 30, ---: 45, ···: 60

The abrupt transitions between the probabilities of signal passage and its stopping at the synapse can be washed out by the stochastic firing of presynaptic neuron-like units. In other words, we have identified another reason to use stochastic elements: The generation of a gradient during learning processes can now be accomplished very natural!

In a network of units there may be other sorts of learning closely related to the mutual interaction of units. Very small weights changes may drive the network into completely new states due to collective phenomena (Haken 1988).

6 Summary

In this contribution we presented an artificial unit with certain similarities to real neurons. Incoming signals were translated into pulse trains in time intervals T . The measurement process only manipulates the firing probability of the unit. Besides the usual delocalisation of information in neural networks through distributed computation, a delocalisation in time was therefore introduced. Similarity between signals was defined via averaging of pulse trains over n periods of T . Thus, the averages carried information, not single pulses.

It turned out that a simple arrangement of logical gates was able to process information in computing roughly a weighted sum of the inputs. A natural consequence was the emergence of a saturation non-linearity which can be encountered also in natural systems.

The weights, realized by stochastically acting gates were identified with synapses of neurons and it was pointed out that the filtering properties of these gates were responsible for the weighting of inputs. Different methods of weight manipulation were outlined. We argued that varying dead-times of gates could lead to a habituation learning which means for synaptic gates to work close to their instability point.

In general, the filtering properties of the synaptic gates can be seen as attractor behavior. Another kind of attractors would appear, if networks from units of the kind described would be combined. Those attractors were comparable to fixed points of other neural net models related to content addressable memory or pattern recognition. Work is now under way to study these effects in more detail.

Acknowledgements. I am indebted to the following persons for very useful comments on parts of this work: Prof. H. Haken, Prof. G. Mahler, K. H. Hasenburg, M. Schindel, C. D. Schulz.

Especially, I want to thank Prof. H. Haken for his continuous encouragement and his support.

References

- Adrian ED (1932) The mechanism of nervous action. University of Pennsylvania Press, Philadelphia
- Bullock, TH, Grinnell A, Orkland R (1977) Introduction to nervous systems. Freeman, San Francisco
- Eccles J (1973) The understanding of the brain. McGraw-Hill, New York
- Feller W (1976) An introduction to probability theory and its applications. Wiley, New York
- Gaines BR (1969) Stochastic computation systems. In: Tou J (ed) Advances in information systems. Plenum Press, New York
- Gaines BR (1987) Uncertainty as a foundation of computational power in neural networks. In: Caudill M, Butler C (eds) Proceedings of the 1st International IEEE Conference on Neural Networks, San Diego, Calif
- Haken H (1988) Private communication
- Hasenburg KH (1987) Private communication
- Hebb DO (1949) The organization of behavior. Wiley, New York
- Hemmerle WJ (1967) Statistical computations on a digital computer. Blaisdell, Waltham, Mass
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci 79:2554–2558
- Lewis ER (1983) The elements of single neurons: a review. IEEE Trans SMC-13:702–710
- McCulloch WS, Pitts W (1943) A logical calculus of ideas immanent in nervous activity. Bull Math Biophys 5:115–133
- Massen R (1977) Stochastische Rechentechnik. Hanser, München
- Omohundro SM (1987) Efficient algorithms with neural network behaviour. Complex Syst 1:273–347
- Peretto P, Niez J (1986) Stochastic dynamics of neural networks. IEEE TA SMC-16:73–83
- Ribeiro ST (1967) Random-pulse machines. IEEE TA EC-16:261–276
- Stein RB (1967) The information capacity of nerve cells using a frequency code, Biophys J 7:797–826
- Steinbuch K (1960) Die Lernmatrix. Kybernetik 1:36–45
- Thompson DW (1917) On growth and form. Cambridge University Press, Cambridge
- Von Neumann J (1956) Probabilistic logics and the synthesis of reliable organisms from unreliable elements. In: Shannon CE, McCarthy J (eds) Automata studies. Princeton University Press, Princeton

Received: July 7, 1988

Dr. W. Banzhaf
 Institut für Theoretische Physik
 und Synergetik der Universität
 Pfaffenwaldring 57/IV
 D-7000 Stuttgart 80
 Federal Republic of Germany