Hebbian Learning in Networks of Spiking Neurons Using Temporal Coding

Berthold Ruf and Michael Schmitt

Institute for Theoretical Computer Science, Technische Universität Graz
Klosterwiesgasse 32/2, A-8010 Graz, Austria
E-mail: {bruf, mschmitt}@igi.tu-graz.ac.at

Abstract. Computational tasks in biological systems that require short response times can be implemented in a straightforward way by networks of spiking neurons that encode analogue values in temporal coding. We investigate the question how spiking neurons can learn on the basis of differences between firing times. In particular, we provide learning rules of the Hebbian type in terms of single spiking events of the pre- and postsynaptic neuron and show that the weights approach some value given by the difference between pre- and postsynaptic firing times with arbitrary high precision. Our learning rules give rise to a straightforward possibility for realizing very fast pattern analysis tasks with spiking neurons.

1 Introduction

Analogue variables within computations in artificial neural networks are usually represented by the average firing rate of some neuron. However, biological neural systems can perform certain computations within a time window being too small for sampling firing rates. It seems more likely that these computations are based on single firing events (see e.g. [1, 10, 13, 12]). For example, Thorpe and Imbert [13] were able to show that humans can analyze and classify visual patterns in 100 msec, although at least 10 synaptic stages are involved. Neurons participating in such computations usually have a firing rate of less than 100 Hz, hence 10 msec are not sufficient to estimate the current firing rate of some neuron. Results like these motivated the interest in investigating networks of spiking neurons, which can base their computation on single firing events.

In this article we consider spiking neuron networks (SNN's), as introduced by Maass [6], where each neuron is basically a *leaky integrate-and-fire* neuron and can be considered as a noise free version of the spike response model by Gerstner and van Hemmen [4]. These SNN's are besides their biological realism also because of their computational power of great interest. In [7, 5] it has been shown that SNN's are computationally more powerful than McCulloch-Pitts neurons (i.e. threshold gates) and also than sigmoidal gates. It has turned out that especially neurons receiving their input as the time difference between firing times can be used to compute in a biologically plausible way the product between its weight vector and its input vector. In [5] it has been shown how this approach can be used for simulating arbitrary feedforward sigmoidal neural nets in a way

which is much faster and more consistent with experimental results about fast information processing in biological neural systems than the usual method using average firing rates. Hence these SNN's can approximate any continuous function of several variables in temporal coding. This approach gives also rise to a very simple and straightforward way of realizing pattern analysis tasks with SNN's [5, 9].

Within this model one considers neurons with given fixed weights that compute functions where the inputs and output are temporally encoded. For example in order to realize pattern analysis one may assume that the learned patterns are stored in weights, whereas new patterns to be analyzed are presented as temporally encoded input. Generally for realizing neural computations based on single firing events it seems to be important to understand the relation between weights and temporally encoded inputs. Thorpe and Gautrais have recently shown that exactly this question could be crucial for fast information processing of neurons in the visual system [12]. They demonstrated that feature extraction might be achieved by the visual system using the order of firing and a proper tuning of synaptic strenghts. Hence the question arises, how a synapse can modify its weight simply on the basis of time differences between firing times and not on firing rates.

Weight changes for some learning process are usually realized according to the Hebb rule, where a synapse is strengthened if both the presynaptic and postsynaptic neuron have at the same time a high firing rate. In this article we want to focus on single firing events, thus we consider some kind of Hebb rule, where the weight change depends on the time difference of single presynaptic and postsynaptic firing times. This type of learning is motivated by recent neurobiological results, where it was shown that the action potential in neocortical pyramidal cells is propagated backwards from the soma to the dendritic tree, such that information about the pre- and postsynaptic firing times is available at the synapse [11] and that synaptic efficacy in those neurons can be actually changed due to this time difference [8].

In this article we focus on supervised learning, where the weight is supposed to assume a certain value and where the weight changes during the learning process are only based on differences between pre- and postsynaptic firing times. This approach is of importance e.g. for the question how a neural system can store information about certain stimuli which are presented repeatedly in the same temporal coding.

We examine how Hebbian learning can be implemented in this context. After introducing the precise model in Section 2 we will show in Section 3 how a single synapse can learn a given weight value simply on the time difference between the presynaptic and the postsynaptic spike. Furthermore we show in Section 4 how several synapses of some neuron can learn given weight values in parallel.

2 Basic Assumptions and Definitions

We consider the common model of *leaky integrate-and-fire neurons*, without noise. As in [5, 6, 7] we make only one additional assumption for realizing computations in such a model: The postsynaptic potential is assumed to grow linearly during an initial segment of fixed but arbitrarily small length.

More precisely, an SNN consists of a finite set V of neurons, a set $E \subseteq V \times V$ of synapses and for each synapse $\langle u, v \rangle \in E$ a response-function $\varepsilon_{u,v}(t) : \mathbf{R}^+ \to \mathbf{R}$ where $\mathbf{R}^+$ denotes the set of all positive reals. One usually models for each neuron v in addition the refractory period of v by a time dependent threshold. However in our constructions, each neuron fires only once within one learning cycle, hence it suffices if each neuron has a constant threshold value Θ. For each synape $\langle u, v \rangle \in E$, where u is the presynaptic and v the postsynaptic neuron, we assume that either $\varepsilon_{u,v}(t) \geq 0$ for all $t \in \mathbf{R}^+$ (excitatory postsynaptic potential, EPSP) or that $\varepsilon_{u,v}(t) \leq 0$ for all $t \in \mathbf{R}^+$ (inhibitory postsynaptic potential, IPSP). The potential $P_v(t)$ of some neuron v at time t is given by

$$P_v(t) = \sum_{u:\langle u,v\rangle \in E} \sum_{s \in F_u : s < t} w_{u,v} \cdot \varepsilon_{u,v}(t-s) \tag{1}$$

where $F_u \subseteq \mathbf{R}^+$ is the set of firing times of neuron u and $w_{u,v} \geq 0$ is the weight of the synapse from neuron u to neuron v. A neuron v fires at time t if $P_v(t)$ reaches Θ from below.

By P_v^{rest} we denote the resting potential of neuron v, i.e. the potential of v when it receives no input. According to (1), $P_v^{\text{rest}} = 0$. However, by continuously receiving sufficiently many EPSP's in addition, P_v^{rest} can be brought arbitrarily close to Θ. (Obviously P_v^{rest} is then no longer time-independent, which corresponds to a noisy membrane potential, however with sufficiently many EPSP's one can reduce this noise arbitrarily.)

Furthermore we assume that each excitatory synapse $\langle u, v \rangle$ has a response function $\varepsilon_{u,v}(t)$ such that

$$\varepsilon_{u,v}(t) = \begin{cases} 0 & \text{for all } t \in [0, d_{u,v}] \\ t - d_{u,v} & \text{for all } t \in [d_{u,v}, d_{u,v} + \delta], \end{cases}$$

where $\delta \geq 0$ is the length of the initial linear segment of the EPSP. The parameter $d_{u,v} \geq 0$ describes the delay between the generation of the presynaptic action potential of neuron u and the time when this action potential reaches the synapse $\langle u, v \rangle$. Throughout this work we consider only passive dendrites, hence we assume that the propagation time in the dendritic tree can be neglected. However even without this assumption one can show similar results as the ones shown in this paper.

3 Monosynaptic Learning

In this section we introduce a Hebbian-type learning rule for monosynaptic learning which is based solely on single pre- and postsynaptic spikes. Let neurons u

and v be connected by an excitatory synapse $\langle u, v\rangle$ with properties described in the previous section. If u fires then a spike arrives at $\langle u, v\rangle$ after a delay of $d_{u,v}$. The synapse does not perceive the actual firing but the arrival of the spike instead. Therefore we will describe firing times of u in terms of arrival times at $\langle u, v\rangle$ in the following.

The learning rule is applied once during a certain period of time which we call learning cycle. The spike trains that are emitted from u and v during such a cycle are shown in Figure 1. The firing of u at t_u is supposed to make v fire at t_v. A second firing of u at t_0 close to t_v is essential for learning in the monosynaptic model in two ways. Firstly, together with t_u it defines a time interval $[t_u, t_0]$ having length $t_0 - t_u$ which can be considered as the input value for v at $\langle u, v\rangle$. Secondly, together with t_v it will be used in the learning rule that depends only on these two firing times during each learning cycle.

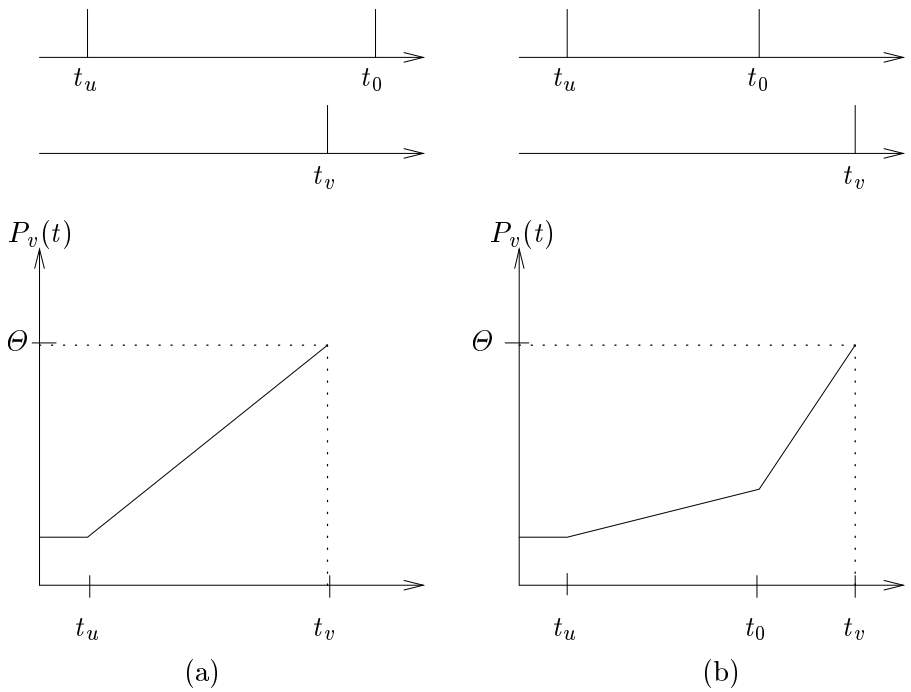

Fig. 1. Spikes at synapse $\langle u, v\rangle$ and potential of neuron v during a learning cycle. In (a) the first spike of u is sufficient to make v fire before the second spike of u; in (b) the firing of v is determined by two spikes of u.

As monosynaptic, Hebbian-type learning rule we propose a change of $w_{u,v}$ by

$$\Delta w_{u,v} \;=\; \eta(t_v - t_0) \tag{2}$$

where $\eta > 0$ is the learning rate, a parameter that has to be chosen before learning starts but is fixed during all learning cycles. The time difference $t_v - t_0$ between the presynaptic spike of u at t_0 and the postsynaptic spike of v at t_v can be considered as an error that has to be reduced during learning.

The main result of this section is a proof that by repeating the learning cycles $w_{u,v}$ approaches a value $\widetilde{w}$ representing the input; in fact, $\widetilde{w}$ is proportional to $1/(t_0 - t_u)$.

Theorem 1. *Given two neurons u, v with excitatory synapse $\langle u, v\rangle$ let the interval $[w_{\min}, w_{\max}] \subseteq \mathbf{R}^+$ be the range of possible values for weight $w_{u,v}$. Let the learning rate η be any arbitrary real number satisfying $0 < \eta \leq w_{\min}^2/(\Theta - P_v^{\text{rest}})$. If neuron u fires twice during each learning cycle with fixed time difference $t_0 - t_u$ and $w_{u,v}$ is updated according to rule (2) with learning rate η then, for any arbitrary initial weight from $[w_{\min}, w_{\max}]$, the sequence of weight values converges to $\widetilde{w}$ such that $\widetilde{w} = (\Theta - P_v^{\text{rest}})/(t_0 - t_u)$.*

Proof. Let $\widetilde{w}$ be the value of $w_{u,v}$ when it remains unchanged. This happens if and only if the second firing of u at t_0 and the firing of v at t_v take place simultaneously, i.e. $t_v = t_0$. Then $\widetilde{w}$ satisfies

$$P_v^{\text{rest}} + \widetilde{w}(t_v - t_u) = \Theta.$$

Substituting t_0 for t_v we obtain

$$t_0 - t_u = \frac{\Theta - P_v^{\text{rest}}}{\widetilde{w}}. \tag{3}$$

Thus, the fixed point $\widetilde{w}$ of the learning rule is proportional to $1/(t_0 - t_u)$.

It remains to show that $\widetilde{w}$ is also an attracting fixed point for all weight values from $[w_{\min}, w_{\max}]$. Let w be the current value of $w_{u,v}$ where $w \neq \widetilde{w}$. We distinguish between two cases, $w > \widetilde{w}$ and $w < \widetilde{w}$. This is necessary because if $w > \widetilde{w}$ then the firing of v at t_v takes place before the second firing of u at t_0. Hence, t_v depends only on $P_v^{\text{rest}}, \Theta$, and t_u. In the case $w < \widetilde{w}$ neuron u fires twice, at t_u and t_0, before $P_v(t)$ reaches Θ. Hence t_v depends also on t_0. We show that in both cases the value of $w_{u,v}$ stays on the same side of $\widetilde{w}$ and approaches $\widetilde{w}$ successively. This will be done using a fixed point theorem from numerical analysis [2].

Case 1: $w > \widetilde{w}$. In this case, which is shown in Figure 1(a), the firing of v at t_v satisfies

$$P_v^{\text{rest}} + w(t_v - t_u) = \Theta$$

from which we get

$$t_v = \frac{\Theta - P_v^{\text{rest}}}{w} + t_u.$$

Using this together with (3) we obtain an expression for the change of w by a single application of rule (2) by

$$\Delta w = \eta(t_v - t_0) = \eta\left(\frac{\Theta - P_v^{\text{rest}}}{w} - \frac{\Theta - P_v^{\text{rest}}}{\widetilde{w}}\right).$$

Thus we can describe the progression of w by an iterative map $f : \mathbf{R} \to \mathbf{R}$ defined as

$$f(w) = w + \kappa\left(\frac{1}{w} - \frac{1}{\widetilde{w}}\right) \quad \text{with } \kappa = \eta(\Theta - P_v^{\text{rest}}).$$

Obviously, $\widetilde{w}$ is a fixed point of f. We will show that f maps $[\widetilde{w}, w_{\max}]$ into itself and is contractive on $[\widetilde{w}, w_{\max}]$. Then it follows by a well known fixed point theorem (see, e.g., [2, Theorem 5.1.1]) that f has a unique fixed point in $[\widetilde{w}, w_{\max}]$, which is then $\widetilde{w}$, and that the sequence $w^{(n+1)} = f(w^{(n)})$ converges to that fixed point for any choice of $w^{(0)} \in [\widetilde{w}, w_{\max}]$.

First we show $f(w) \geq \widetilde{w}$ for all $w \in [\widetilde{w}, w_{\max}]$. This is equivalent to

$$w + \kappa/w \geq \widetilde{w} + \kappa/\widetilde{w} \quad \text{for all } w \in [\widetilde{w}, w_{\max}]. \tag{4}$$

The condition for the learning rate $\eta \leq w_{\min}^2/(\Theta - P_v^{\text{rest}})$ can be rewritten as $w_{\min}^2 \geq \kappa$ from which

$$1 - \frac{\kappa}{w^2} \geq 0 \quad \text{for all } w \geq w_{\min} \tag{5}$$

follows. The expression $1 - (\kappa/w^2)$ is the derivative of $g(w) = w + (\kappa/w)$ which is therefore increasing on $[w_{\min}, w_{\max}]$. Hence (4) is proven and $f(w) \geq \widetilde{w}$ holds for all $w \in [\widetilde{w}, w_{\max}]$.

From $w > \widetilde{w}$ we get

$$\kappa\left(\frac{1}{w} - \frac{1}{\widetilde{w}}\right) < 0.$$

Thus $f(w) < w$ holds which implies $f(w) < w_{\max}$ for all $w \in [\widetilde{w}, w_{\max}]$. It follows from this, together with $f(w) \geq \widetilde{w}$, that f maps $[\widetilde{w}, w_{\max}]$ into itself.

To see that f is contractive on $[\widetilde{w}, w_{\max}]$ we consider f', the derivative of f, and show that $|f'(w)| \leq L$ for all $[\widetilde{w}, w_{\max}]$ with Lipschitz constant $L < 1$. The derivative of f is

$$f'(w) = 1 - \frac{\kappa}{w^2}.$$

In (5) this expression was already seen to be non-negative. A Lipschitz constant $L < 1$ such that $f'(w) \leq L$ for all $w \in [\widetilde{w}, w_{\max}]$ is then provided by $L = 1 - (\kappa/w_{\max}^2)$.

Case 2: $w < \widetilde{w}$. Here the potential of neuron v is determined by two spikes of neuron u, at t_u and at t_0 (see Figure 1(b)). The firing time t_v of v then satisfies

$$P_v^{\text{rest}} + w(t_v - t_u) + w(t_v - t_0) = \Theta$$

from which we derive

$$t_v = \frac{1}{2}\left(\frac{\Theta - P_v^{\text{rest}}}{w} + t_0 + t_u\right).$$

Solving (3) for t_0 and substituting we can rewrite this as

$$t_v = \frac{1}{2}\left(\frac{\Theta - P_v^{\text{rest}}}{w} + \frac{\Theta - P_v^{\text{rest}}}{\widetilde{w}}\right) + t_u.$$

Combined with (3) this yields a synaptic change of

$$\Delta w = \eta(t_v - t_0) = \frac{\eta}{2}\left(\frac{\Theta - P_v^{\text{rest}}}{w} - \frac{\Theta - P_v^{\text{rest}}}{\widetilde{w}}\right).$$

The sequence of weights is then determined by iterating the function $h : \mathbf{R} \to \mathbf{R}$ defined as

$$h(w) = w + \lambda\left(\frac{1}{w} - \frac{1}{\widetilde{w}}\right) \quad \text{with } \lambda = \eta(\Theta - P_v^{\text{rest}})/2.$$

Obviously, $\widetilde{w}$ is a fixed point of h. It remains to show that h maps $[w_{\min}, \widetilde{w}]$ into itself and is contractive on $[w_{\min}, \widetilde{w}]$. From $\lambda = \kappa/2$ and (5) we get

$$w + \lambda/w \leq \widetilde{w} + \lambda/\widetilde{w} \quad \text{for all } w \in [w_{\min}, \widetilde{w}]$$

which is equivalent to $h(w) \leq \widetilde{w}$ for all $w \in [w_{\min}, \widetilde{w}]$. From $w < \widetilde{w}$ we get $\lambda(1/w - 1/\widetilde{w}) > 0$ and hence $h(w) > w_{\min}$ for all $w \in [w_{\min}, \widetilde{w}]$. Thus h maps $[w_{\min}, \widetilde{w}]$ into itself. Finally, along the same lines of reasoning as in case 1, it is now easy to conclude that h is contractive on $[w_{\min}, \widetilde{w}]$ with Lipschitz constant $L = 1 - (\lambda/w_{\max}^2) < 1$.

This completes the proof of the theorem. □

With regard to the convergence rate we can make the following observation: It is well known that for contractive iterative mappings the rate of convergence is governed by

$$|w^{(n)} - \widetilde{w}| \leq L^n |w^{(0)} - \widetilde{w}|$$

where $L < 1$ is a Lipschitz constant [2]. Thus the sequence of weights generated according to the learning rule (2) satisfies

$$|w^{(n)} - \widetilde{w}| \leq \left(1 - \frac{\mu}{w_{\max}^2}\right)^n |w^{(0)} - \widetilde{w}|$$

where $\mu = \eta(\Theta - P_v^{\text{rest}})$ if $w^{(0)} > \widetilde{w}$ and $\mu = \eta(\Theta - P_v^{\text{rest}})/2$ if $w^{(0)} < \widetilde{w}$.

In the theorem we put no restrictions on values $w_{\min}$ and $w_{\max}$, except that they have to be positive, in order to keep the statement and the proof as general as possible. However, for learning to be biologically plausible, restrictions are certainly necessary. For instance, it seems adequate that the pre- and postsynaptic firing times t_0 and t_v, on which the weight change is based, are not too far from each other. On the other hand, our construction relies on the fact that the pair of pre- and postsynaptic spikes at t_u and t_v does not result in a change. Hence they should not be too close to each other. Furthermore, it must be guaranteed that the firing of v caused by u takes place while the EPSP is still in its initial linear segment. By raising the resting potential P_v^{rest} as described in Section 2

and by restricting the range of possible weight values these requirements can be met easily.

The type of learning we have considered in this section is based on three spikes during each learning cycle, two presynaptic and one postsynaptic. The pair of presynaptic spikes is used to describe the input value, whereas one presynaptic and the postsynaptic spike define the synaptic modification. In the context of supervised learning this seems to be the minimal number of spikes required for biologically plausible, Hebbian-type learning.

4 Parallel Learning

The method described in the previous section makes it possible to learn single weights. However, usually a neuron v receives its input from several neurons $u_1, \ldots, u_n$. The corresponding weights $w_{u_1,v}, \ldots, w_{u_n,v}$ can still be learned in the above described way, if the weights are learned sequentially, such that v receives only inputs through one synapse during one learning cycle. Hence the question arises, how several weights can be learned in parallel without too strong additional assumptions. If one would try to learn weights in parallel in a similar fashion as described in the previous section, the problem would arise of how to distribute the error encoded by $t_0 - t_v$ among the weights. Additional error signals would make Hebbian learning, where the weight change is only influenced by the time difference between some pre- and postsynaptic firing time, impossible.

Since the traditional formulation of the Hebb-rule allows only an increase of the weights, one frequently assumes in addition that after each weight change all weights of the neuron are normalized (see e.g. [3]). Hence if one normalizes the weight vector $\underline{w} = (w_{u_1,v}, \ldots, w_{u_n,v})$ such that $||\underline{w}|| = 1$ then a simple Hebbian learning rule can be formulated: The goal is that $\underline{w} = \underline{\widetilde{w}}$ for some desired weight vector $\underline{\widetilde{w}}$ with $||\underline{\widetilde{w}}|| = 1$. In contrast to Section 3 v receives here through some additional synapse with a sufficiently strong weight an input spike such that v fires at time t_0. The i-th synapse $\langle u_i, v\rangle$ receives a spike from neuron u_i at time t_{u_i} such that $t_0 - t_{u_i} = \widetilde{w}_{u_i,v}$. We ensure that v does not fire before t_0 by providing sufficiently many IPSP's through additional inhibitory synapses. One gets the learning rule

$$\Delta w_{u_i,v} = \eta(t_0 - t_{u_i}) \,/\, ||\underline{w}||\,, \qquad 1 \leq i \leq n, \tag{6}$$

where $\eta > 0$ is the learning rate, that has the following convergence property:

Theorem 2. *Let $\underline{\widetilde{w}}$ be a weight vector satisfying $||\underline{\widetilde{w}}|| = 1$ and let $\eta > 0$. Then the sequence of weight vectors generated by the parallel learning rule (6) converges to $\underline{\widetilde{w}}$ for any arbitrary initial weight vector.*

Proof. We immediately have $\underline{w}^{(n)} \cdot \underline{\widetilde{w}} \leq 1$ for $n \geq 1$ by the Cauchy-Schwarz inequality. It is not hard to see then that $\underline{w}^{(n)} \cdot \underline{\widetilde{w}} < \underline{w}^{(n+1)} \cdot \underline{\widetilde{w}}$ holds for $\underline{w}^{(n)} \neq \underline{\widetilde{w}}$, which implies the convergence. □

It is easy to verify that by some proper transformation of the normalization not only weights from $[0,1]$ but from some arbitrary interval $[w_{\min}, w_{\max}] \subseteq \mathbf{R}^+$ can be learned.

We finally observe that the type of learning in this section is not based on an error signal. Furthermore, we do not require here that EPSP's have a linear segment.

5 Conclusions

We have introduced some type of Hebbian learning for spiking neurons by providing rules for synaptic modification that are based on single firing events of the pre- and postsynaptic neuron. To the best of our knowledge, this is the first approach for this model where it is shown that the weights provably converge to the desired values. Thus, a spiking neuron can learn weights from its firing times, which appears to be essential for fast information processing in biological neural systems. Furthermore, our constructions give rise to simple methods for learning weights within pattern analysis tasks where the inputs are temporally encoded.

In the model of a spiking neuron we have explicitly excluded the presence of noise. However, our constructions in Sections 3 and 4 should also work in the case of a certain amount of noise, but they are harder to analyze theoretically. We expect that similar results can be shown using standard methods dealing with noise.

References

1. Abeles, M., Bergman, H., Margalit, E., and Vaadia, E.: Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. J. of Neurophysiology **70** (1993) 1629–1638.
2. Blum, E.K.: Numerical Analysis and Computation: Theory and Practice. Addison-Wesley, Reading, Mass. (1972).
3. Brown, T.H. and Chattarji, S.: Hebbian synaptic plasticity. In: Handbook of brain theory and neural networks, Arbib, M., ed., MIT Press, Cambridge, (1995) 454–459.
4. Gerstner, W. and van Hemmen, L.H.: How to describe neuronal activity: spikes, rates, or assemblies? Advances in Neural Information Processing Systems 6, Morgan Kaufmann, San Mateo (1994) 463–470.
5. Maass, W.: Fast sigmoidal networks via spiking neurons. To appear in Neural Computation.
6. Maass, W.: Lower bounds for the computational power of networks of spiking neurons. Neural Computation **8** (1996) 1–40.
7. Maass, W.: Networks of spiking neurons: the third generation of neural network models. Proc. of the 7th Australian Conference on Neural Networks, Canberra, Australia (1996) 1–10.
8. Markram, H.: Neocortical pyramidal neurons scale the efficacy of synaptic input according to arrival time: a proposed selection principle of the most appropriate synaptic information. Proc. of Cortical Dynamics in Jerusalem (1995) 10–11.

9. Ruf, B.: Pattern analysis with networks of spiking neurons. Submitted for publication.
10. Sejnowski, T. J.: Time for a new neural code? Nature **376** (1995) 21–22.
11. Stuart, G. J. and Sakmann, B.: Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. Nature **367** (1994) 69–72.
12. Thorpe, S. J. and Gautrais, J.: Rapid visual processing using spike asynchrony. To appear in Advances in Neural Information Processing Systems 9, MIT Press, Cambridge, 1997.
13. Thorpe, S. J. and Imbert, M.: Biological constraints on connectionist modelling. In: Connectionism in Perspective, Pfeifer, R., Schreter, Z., Fogelman-Soulié, F., and Steels, L., eds., Elsevier, North-Holland (1989).

This article was processed using the LaTeX macro package with LLNCS style