

Learning at the edge of chaos : Temporal Coupling of Spiking Neurons Controller for Autonomous Robotic

Hédi Soula and Aravind Alwan and Guillaume Beslon

ALAB Team of Prisma Lab
Computer Science Dept National Institute of Applied Science
LYON – FRANCE

Abstract

In this paper, a recurrent spiking neural networks is trained on an robot to learn to avoid obstacles using visual flow. At the starting of the process, this network is initialized in a "chaotic" state and a STDP-like learning algorithm is used. We argue that a proper scaling variable can direct the network from chaos to synchronized state and back. This process allows us to train the robot because it links (external) temporal loops with (internal) neural activity. We use the scaling factor to have this coupling functional. Given an over-simplistic scaling, we managed to obtain very interesting resulting behaviors when tested on a real robot.

Introduction

The problem of adaptation of robots in an environment is a challenging issue. The obvious problem is the balance between real adaptation and memory. In other words, an adaptive device must be both plastic (in order to learn new things) and rigid (not to forget them).

Moreover, a behavior is a spatio-temporally structured relationship between an robot and its environment. Thus it is a temporal process where past has a consequence on the future. A good adaptation mechanism should also allow a temporal retrieval of sequence of elementary behaviors in an efficient way. As a consequence, adaptation must be viewed as a result from coupling the dynamics of input/output time series with dynamics inside an artificial brain on a codependent way.

However, even from that point of view, the problem still needs to be solved so far. Many architectures proposed dynamical process for temporal series learning. Most of the time, the learning procedures are off-line and supervised. An alternative was found in genetic algorithm (Soula, Beslon, & Favrel 2003; di Paolo 2003; Floreano & Mattiussi 2001). Unfortunately, all these approaches still lack of on-line adaption methods and offer no obvious path for enabling them.

Still, some approaches stands on the ground of dynamical systems in a way to make an original use of

their properties. Indeed, seen beforehand as a problem to avoid, the chaotic nature of certain architecture appears as a new paradigm for learning. In these models memory is not seen anymore as a fixed point to where the process should converge.

These facts will be our starting point. At the start, the controller must exhibit enough intrinsic features to encompass wide range of dynamics – a chaotic state. Adaptation will therefore be a plasticity mechanism that allows us to put constraints on "good" dynamics. On the other hand, it must also have the power to "go back" into a general disordered state (Langton 1991).

Throughout this paper, we will argue that recurrent spiking neuron networks have the required properties. The mechanism of plasticity is inspired from biology : the Spike-Timing Dependent Plasticity. We show the effect of such rules on the overall behavior of a recurrent spiking neural network.

Finally, we propose a method using the STDP-rule for learning behaviors. Indeed, using an over-simplistic scaling, we show the results of our experiments to train a real robot to avoid obstacles using its visual flow.

The model

Temporal Coding

The biological brain rests on a network of neurons which communicate by sending pulses one with each other via their axons. These pulses (also known as spikes) are small current charges emitted when the membrane potential of the neuron crosses a threshold. The membrane potential is then reset and begins to grow again according to the input spikes coming from other neurons. Moreover, for a short period after spiking - called refractory period, the neuron can no longer spike no matter how strong is the input stimulation.

Supported by many neurobiological observation (e.g (Adrian 1926)), major contributions to artificial neural networks came from the frequency-based models of such neurons. However, contemporary neurobiology suggests that a smaller temporal scale (enabling to use the exact timing of a spike) is of major importance to understand neural treatment. Indeed, determination of temporal and spatial correlations in neural activity has

become one of the most important tools in neurophysiological brain research. During the last decade, temporal synchronizations between far-off neurons have been detected in most of the cerebral areas (Gray *et al.* 1989; Singer 1993; Varela *et al.* 2001) as well as strongly stimulus dependent correlations (Gray 1999).

On another point of view, the process time would be too long if at each step neurons should wait and integrate for a long time in order to convey a firing rate. The example of vision recognition is particularly striking (Thorpe, Fize, & Marlot 1996). For example, (Hopfield 1995; Samuelides, Thorpe, & Veneau 1997) have shown that the exact firing date of a neuron can precisely code the nature of stimulus received by this neuron. It is then useless to wait for more spikes to transmit the neural information to the remaining of the net (Maas & Bishop 2001; Gerstner & Kistler 2002).

Finally, from a theoretical point of view, many authors stress that the spike dynamics of a network of neurons is of central importance to elucidate components of the neural code in cortex - to understand the mechanisms of cross-correlations (Meyer & van Vreeswijk 2002) and the temporally irregular firing in the cortex (van Vreeswijk & Sompolinsky 1996).

Spiking neurons

In this section, we present the model of discrete leaky integrate and fire neuron (Lapicque 1907; Tuckwell 1988; Domany, Hemmen, & Schulten 1994) and the notations that will be used throughout the remaining of the paper.

In this model, the neuron state is defined by one state variable : the membrane potential V . The dynamics of the neuron is defined by the following discrete equation :

$$V_i(t) = \gamma(V_i(t-1) - V_{rest}) + \sum_{j=0}^N W_{ij} \delta(t - T_j - d_{ij}) + I_i(t)$$

where i and j are indexes of neurons (all the neurons j are inputs of i - N pre-synaptic neurons in all). γ is a decay rate $\gamma \in [0, 1]$. V_{rest} is the potential of rest (and reset). $I_i(t)$ is an external input current for neuron i . W_{ij} are the synaptic strengths (the weights). T_j is the last time neuron j fired, d_{ij} is an axonal delay between j and i . Finally δ is the Kronecker function.

When V_i crosses the threshold $\theta(t)$, the neuron fires and is reset to its resting potential. We can express this transition:

$$\begin{aligned} \lim_{\Delta t \rightarrow 0^+} V_i(T_i - \Delta t) &= \theta(t) \\ \lim_{\Delta t \rightarrow 0^+} V_i(T_i + \Delta t) &= V_{rest} \end{aligned}$$

$\theta(t)$ is a normal random variable with mean $\bar{\theta}$ and variance σ_{noise}^2 . A random threshold introduces noise in the spiking dates and consequently in all the network.

Setting $T_i^0 = -\infty$ we define T_i^n (the n -th spiking date) recursively as this:

$$T_i^n = \inf(t \mid t > T_i^{n-1} + r_i, V_i(t) \geq \theta(t))$$

where r_i is the refractory period for the neuron i .

Leaky integrate and fire neurons are known to be good approximates of biological neurons concerning spiking dates distribution. Moreover, they are simple enough and easy to handle when embedded in a robot. For the all the experiments, $V_{rest} = 0.0$, $\bar{\theta} = 1.0$, $r_i = 4$ for all i and $\gamma = 0.99$.

Dynamics of spiking-neurons nets

As a general approach, some authors (Daucé 2000; Daucé & Quoy 2000; Maas, Natschlager, & Markram 2002) used the dynamical properties and chaotic features of recurrent networks of classical sigmoid neurons as a tool for learning and pattern storing. In the spiking neurons case, it is still to be done.

At a first glance, spiking neural networks show a very broad variety of dynamics (Golomb 1994; van Vreeswijk & Sompolinsky 1996; Meyer & van Vreeswijk 2002). In simple case enough it was determined. Sufficient condition for phase synchronization and its stability was proposed in homogeneous networks (Chow 1998; Gerstner 2001). Chaotic properties (expressed as Lyapunov exponents) was discovered for particular symmetric networks (Coombes 1999). In the precise case of integrate and fire neurons, equilibrium criteria has been calculated for nets of irregular firing neurons (Brunel 2000; Amit & Brunel 1997b; 1997a) and VLSI neurons (Fusi & Mattia 1999; Mattia & Giudice 2000).

In a behavioral perspective (i.e. the spiking neuron network will be in charge of a dynamic sensori-motor task), it becomes vital to examine the controller internal properties regarding its connectivity. More precisely, the network will have to exhibit two contradictory dynamic features. On one hand, It must be reactive - input stimuli should modify the dynamics. On the other hand, when available information in a noisy environment consists on coherent flow of inputs, it must have its own internal dynamics to be able to maintain a stable behavior.

Any internal dynamics emerge from the collective behavior of interacting neurons. This is a product of the coupling between neurons. So to start the study, we begin by examine the influence of a coupling factor - the intensity of the average influence of one neuron upon one another.

Assume we have a stochastic totally recurrent spiking neural nets. Suppose the weight distribution follows a centered normal law. The coupling factor here is the variance of the law σ_w^2 . In order to check the effect of the variance we introduce the average potential of the network :

$$m(t) = \frac{1}{N} \sum_{i=1}^N V_i(t)$$

Figures 1 and 2 show the temporal evolution of the average potential when the network receives a unique stimulation at the start of the process (so-called spontaneous mode). In order for the regime to be self-sustaining (i.e. no neural death) the coupling factor

must be high enough. When it is so, the dynamics evolves from a chaotic mode (in figure 1) to a synchronous mode (in figure 2).

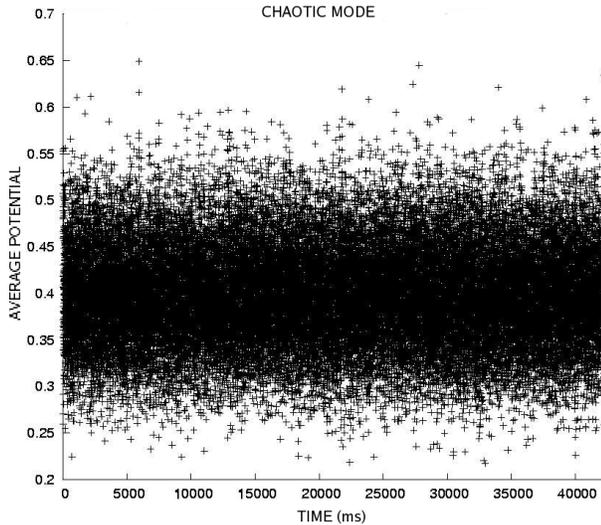


Figure 1: Chaotic mode. Parameters : $N = 100$, $r = 4$ and $\sigma_w = 0.18$

The chaotic mode expresses that the interactions between neurons are too weak to stick to some predetermined pattern. However, with high coupling the dynamics is “locked” in a very stable state.

At the neuron level, in the chaotic mode, neurons fire asynchronously and aperiodically. Increasing value of variance leads to an increase of periodicity among neurons as well as synchrony.

When used as a controller, the internal mode imposes different behavior to an agent. A network in chaotic coupling gives a agent with input-led and almost random behavior. This is a particularly interesting feature for exploration problem. On the other hand, synchronous mode gives a stereo-typical (“autistic”) behavior for the agent. In this case, the internal dynamics dominates the flow of input (leading to exploitation).

Obviously, these two important behaviors must be combined. Thus the “promised land” must be somewhere in between – “at the edge of chaos”. Starting from a chaotic state, a recurrent network while learning should be able to evolve toward both types of functioning. It should evolve toward more synchrony to manage a coupling between the environment and the agent and “back” to chaos to test acquire new (and possibly better) behaviors.

Learning at the Edge of Chaos

Recent neurobiology experiments have suggested that the relative timing of pre- and post-synaptic potentials played an important role in determining the intensity as well as the sign of variation of a synapse strength

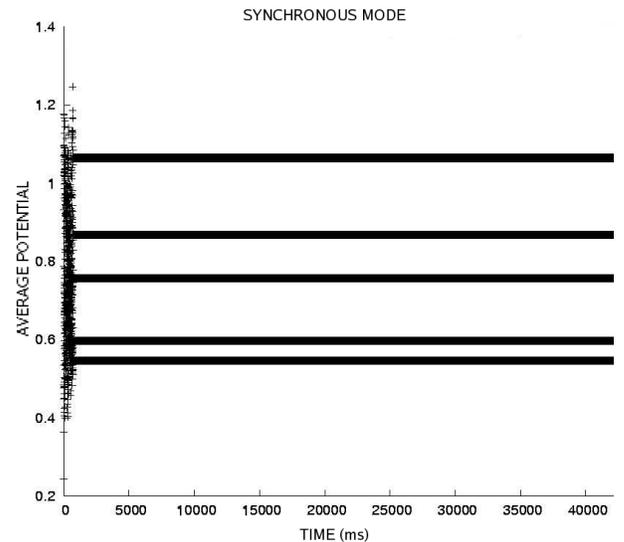


Figure 2: Synchronous mode. Parameters : $N = 100$, $r = 4$ and $\sigma_w = 0.30$

(Markram *et al.* 1997; Bi & Poo 1998; Abbott & Nelson 2000). The intensity of this Long Term Potentiation (and Depression) is directly dependent of the relative timing – the spike delay between the post-synaptic and pre-synaptic neurons. In addition, if this delay is high enough (order of tens of milliseconds) no modification occurs. On the other hand, the modification is maximal when the post-synaptic neuron fires just after (or just before) the pre-synaptic does.

As (Bi & Poo 1998) put it, one can extract quite straightforwardly a very simple rule that rests upon inter-spikes delays. This “rule” is known as Spike-Time Dependent Plasticity (STDP). It became a widespread implementation of Hebb’s initial intuition on memory formation in the brain (Hebb 1949).

STDP rule involves weight modification as this :

$$\Delta W_{ij} = \alpha(1 - |W_{ij}|)h_d(\Delta_{ij})$$

where Δ_{ij} is the difference between the last firing dates of post-synaptic neuron i and pre-synaptic j . $h_d(t)$ is a function which depends on the axonal delay between j and i . α is a scaling parameter. We chose the simplest possible shape for h_d : piecewise linear and anti-symmetric defined as :

$$h_d(t) = \begin{cases} \frac{T-t}{T-d} & d < t < T \\ \frac{t}{d} & 0 < t < d \\ 0 & t > T \\ -h_d(-t) & t < 0 \end{cases}$$

Here T is the time-out constant (i.e the relative timing above which no modification occurs).

The STDP rule was chosen because it implies synchronization between neurons. At the start, we get a network in a chaotic spontaneous state. Then, the

STDP rules is applied with a positive scaling factor. The temporal evolution of the average potential is showed in figure 3. During the process, the characteristic cloud of chaos begins to leave room for no-less characteristic stripes of the synchronous mode. When inverting the sign (anti-STDP), the network is moving back from synchronous mode toward chaotic dynamics (from stripes to cloud).

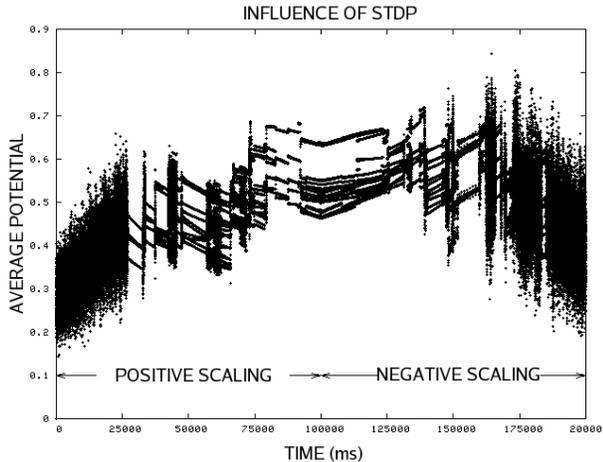


Figure 3: Influence of STDP algorithm on the coupling. It shows the evolution of the Mean Potential over the time. During the first 100000 ms, the scaling factor α is > 0 and is < 0 for the remaining time. We start with a chaotic network. Parameters: $N = 100$, $\sigma = 0.2$, $r = 4$, $T = 50$, $|\alpha| = 0.05$.

As such the STDP/anti-STDP algorithm will be our learning framework. Starting with a random network in a chaotic mode, an agent will freely explore. The sign of the scaling factor will be adjusted in order to synchronize/desynchronize when needed allowing the network to “learn at the edge of chaos”.

Experiments and results

We tested our approach on a task of obstacle avoidance with visual flow. The robot has to avoid walls and obstacles using only its camera information. More precisely, there will be neither proximity sensors nor positioning device available to it. In addition, the environment layout will not allow the robot to extract any simple rule to compute its exact position. Obviously in order to accomplish such a difficult task, the network must exhibit important internal loops since no static input provides by itself enough information.

Thus, we used a Khepera robot equipped with the linear *k213* camera turret. It is positioned in an arena with black and white vertical stripes of random size painted on the walls at irregular intervals. Moreover, up to three black cylinder-shaped obstacles are scattered around. It is a similar environment than described in (Floreano & Mattiussi 2001) (see figure 4).

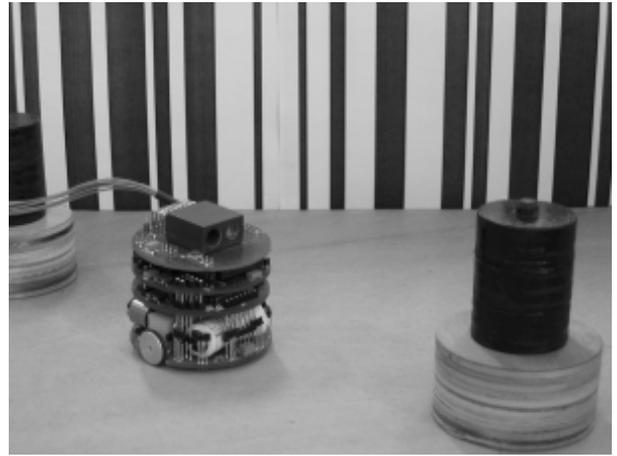


Figure 4: The experiment arena.

The controller of the robot is a spiking neural network with three layers of neurons. The first and third serve as sensors and motors neurons respectively.

More precisely, the 64 pixels of the *k213* linear camera are averaged to provide an array of 16 values which correspond to the 16 input pixel neurons. These neurons are fed with input current allowing frequencies between 10Hz and 100Hz. The current value I is calculated to provide the neuron a desired period P . It was done simply using the formula :

$$I = \theta \frac{1 - \gamma}{1 - \gamma^P}$$

We recall that γ is the leak (decay rate) and θ the threshold. Both were constant throughout experiments and identical for all neurons (Parameters : $\gamma = 0.99$ and $\theta = 1.0$). Hence, the values (in grey level) of the camera pixel are scaled between two frequencies (10Hz for full black pixels and 100Hz for full white pixels).

The 2 output neurons serve as motoneurons – one for each motor (left and right). The motor speeds are proportional to the corresponding neuron firing rate.

The intermediate layer – the hidden layer – consists on all-to-all connected 40 neurons. Each input neuron has a connection to each hidden neurons. The weights of this connection are chosen randomly using a centered normal law. The standard deviation for the inputs weights is 0.5. The hidden layer is completely recurrent and as argued in the previous sections, the standard deviation for the weights is set to obtain a chaotic network in spontaneous mode (0.2 for 40 neurons). Finally all hidden layer neurons have a connection to the two motoneurons. The standard deviation is 0.2.

Between two input informations accessing (approximately 100 time steps), we let the network evolves freely.

At each time step, STDP/anti-STDP learning is applied with scaling factor $\alpha = \omega R$. ω is constant during the whole experiment ($\omega = 0.005$) and R follows the

rule below :

- $R = 1.0$ if the robot moves forward,
- $R = -1.0$ if the robot hits a wall or an obstacle,
- $R = 0.0$ otherwise.

In other words, when the robot moves forward, we try to synchronize the internal dynamics with the external evolving visual flow. We did it in the opposite sense when it hits an obstacle. In the rest of the time that is when it stops, move rearward or around, we do nothing.

At the start of a trial, a robot will move randomly and explore its environment. It will by accident hit an obstacle or move forward. We tested 10 different random network during 150 s (in simulated neural time). We counted the number of shocks and the number of time the robot moved forward. The results for a typical individual are displayed in figure 5.

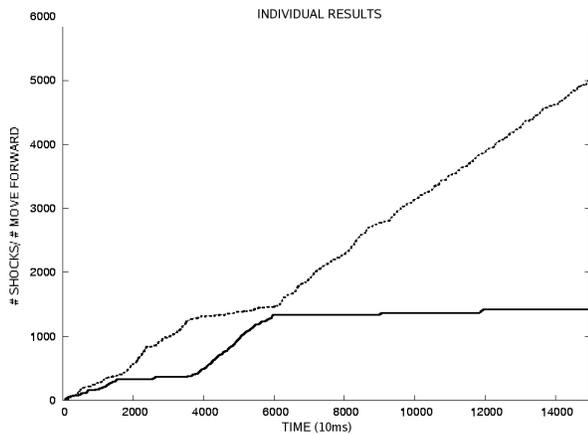


Figure 5: Number of shocks (plain line below) and number of steps forward (dashed line above)

The obstacle avoidance was learned for 9 of the 10 networks tested. Even if some of them were very performant the avoidance was never robust. Since the task need a complete (and fine) coupling between the input flow and its own speed, the robot was able to stop and turn only when the obstacles appeared at high speed (relative to the robot). We tried to move by hand the obstacles in front of the robot. The obstacle avoidance is observed when obstacle is moved quickly. This indicates that the internal dynamics is able to extract the speed of the visual flow (which, here, is the only way to avoid the obstacles). However, if the obstacle approaches slowly, the robot is often not able to avoid it.

Conclusion and perspectives

As argued throughout this article, the path toward total adaptiveness (in the sense of adapting its own behavior and behavioral structure) rests on the principles of co-dependence between the robot and its environment.

Respecting these principles enables us to develop and study robots' behaviors as well as their structures. It is a key approach to learning.

Moreover, the critical stage lies in the intrinsic feature of the controller – a dynamical controller for a dynamical environment. Thus, for co-dependence to emerge one needs for the controller to exhibit various dynamics. In this case, the chaotic nature of recurrent spiking neural network is a very interesting feature. Indeed, “learning at the edge of chaos” is a powerful way to assure an emerging coupling between external and internal dynamics.

However, we aware of the few steps made in this direction. Indeed, empirical use of this Hebbian rule may not be enough to extract more than simple behaviors. Orienting the learning toward an observed behavior corresponding with our will will probably be a much more complicated task. Still, it is obviously the prospect for future development and work.

Indeed, we have to keep in mind that, in our simple experiments, the results were considerably more fruitful concerning the structuration of behaviors - via notably synchronization - than, for example, genetic adaptation of weights. This validates partly our approach even if a deeper understanding of the results is obviously needed.

References

- Abbott, L. F., and Nelson, S. B. 2000. Synaptic plasticity: taming the beast. *Nature America* 3:1178–1182.
- Adrian, E. D. 1926. The impulses produced by sensory nerve endings. *J. Physiol.* (61):49–72.
- Amit, D., and Brunel, N. 1997a. Dynamics of recurrent spiking neurons before and following learning. *Network: Comput. Neural. Syst.* 8:373–404.
- Amit, D., and Brunel, N. 1997b. Model of global spontaneous activity and local structured delay activity during learning periods in the cerebral cortex. *Cerebral Cortex* 7:237–252.
- Bi, G., and Poo, M. 1998. Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience* 18(24):10464–10472.
- Brunel, N. 2000. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience* 8:183–208.
- Chow, C. 1998. Phase-locking in weakly heterogeneous neural networks. *Physica D*.
- Coombes, S. 1999. Chaos in integrate-and-fire dynamical systems. In *In proc. of Stochastic and Chaotic Dynamics in the Lakes*. American Institute of Physics Conference Proceedings.
- Daucé, E., and Quoy, M. 2000. Random recurrent neural networks for autonomous system design. In *proc. of the sixth Int. Conf. on Simulation of Adaptive Behavior*.

- Daucé, E. 2000. *Adaptation Dynamique et Apprentissage dans les Réseaux de Neurones Récurrents Aléatoire*. Ph.D. Dissertation, ENSAE.
- di Paolo, E. 2003. Evolving spike-timing-dependent plasticity for single-trial learning in robots. *Phil. Trans. R. Soc. Lond. A* 361:2299–2319.
- Domany, E.; Hemmen, J. V.; and Schulten, K., eds. 1994. *Model of Neural Networks II Temporal Aspects of Coding and Information Processing in Biological Systems*. Springer-Verlag.
- Floreano, D., and Mattiussi, C. 2001. Evolution of spiking neural controllers for autonomous vision-based robots. In Gomi, T., ed., *Evolutionary Robotics*. Berlin: Springer-Verlag.
- Fusi, S., and Mattia, M. 1999. Collective behavior of networks with linear (vlsi) integrate-and-fire neurons. *Neural Computation* 11:633–652.
- Gerstner, W., and Kistler, W. 2002. *Spiking Neuron Models. Single Neurons, Populations, Plasticity*. Cambridge University Press.
- Gerstner, W. 2001. Populations of spiking neurons. In Maas, W., and Bishop, C., eds., *Pulsed Neural Networks*.
- Golomb, D. 1994. Clustering in globally coupled inhibitory neurons. *Physica D* 72:259–282.
- Gray, C.; Köning, P.; Engel, A.; and Singer, W. 1989. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338(6213):334–337.
- Gray, C. 1999. The temporal correlation hypothesis of visual feature integration : Still alive and well. *Neuron* 24:31–47.
- Hebb, D. 1949. *The Organization of Behavior*. Wiley New York.
- Hopfield, J. 1995. Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376(6535):33–36.
- Langton, C. 1991. Computation at the edge of chaos: Phase transition and emergent computation. In *Emergent Computation*. The MIT Press.
- Lapicque, L. 1907. Recherche quantitatives sur l'excitation électrique des nerfs traitée comme polarisation. *J. Physiol Pathol. Gen.* 9:620–635.
- Maas, W., and Bishop, C., eds. 2001. *Pulsed Neural Networks*. Cambridge, Massachusetts: MIT Press.
- Maas, W.; Natschlag, T.; and Markram, H. 2002. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation* 14:2531–2560.
- Markram, H.; Lubke, J.; Frotscher, M.; and Sakmann, B. 1997. Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science* 275:213–215.
- Mattia, M., and Giudice, P. D. 2000. Population dynamics of interacting spiking neurons. *Physical Review E* 66(5).
- Meyer, C., and van Vreeswijk, C. 2002. Temporal correlations in stochastic networks of spiking neurons. *Neural Computation* 14(2):369–404.
- Samuelides, M.; Thorpe, S.; and Veneau, E. 1997. Implementing hebbian learning in a rank-based neural network. In *Proc. of ICANN'97, Lausanne*, 145–150.
- Singer, W. 1993. Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology* (55):349–374.
- Soula, H.; Beslon, G.; and Favrel, J. 2003. Evolving spiking neurons nets to control an animat. In Pearson, D.; Steel, N.; and Albrecht, R., eds., *Proc. of ICANN-GA 2003, Roannes*, 193–197.
- Thorpe, S.; Fize, D.; and Marlot, C. 1996. Speed of processing in the human visual system. *Nature* 381:520–522.
- Tuckwell, H. 1988. *Introduction to theoretical neurobiology: Vol.2:Non linear and stochastic theories*. Cambridge, Massachusetts: Cambridge University Press.
- van Vreeswijk, C., and Sompolinsky, H. 1996. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274:1724–1726.
- Varela, F.; Lachaux, J.-P.; Rodriguez, E.; and Martinerie, J. 2001. The brainweb: Phase synchronization an large-scale integration. *Nature Neuroscience Review* 2:229–239.