

Coherence, Memory and Conditioning

A Modern Viewpoint

Raul C. Mureşan^{1,2}

1. Nivis Research

Gh. Bilaşcu 85

400489 Cluj-Napoca, Romania

2. Faculty of Automation and Computers

Technical University of Cluj Napoca

26-28 Baritiu Street

400027 Cluj-Napoca, Romania

email: raulmuresan@yahoo.com

<http://www.raulmuresan.home.ro>

Abstract

We present a novel viewpoint on the complexity of neural mechanisms, addressing some aspects of cortical processing, like memory and modulation and coherence. This study shows how relatively small, recurrent, microcircuits can interact with populations of neurons, achieving memory effects with spontaneous recall and how context modulation could play an important part in conditioned recall.

The emphasis is on the dynamical self-sustained activity of the recurrent microcircuit viewed as a "mental state" which can be interpreted by a target output population and controlled by external waves of activity (mainly external inhibitory waves). We also address the problem of self-organization and dynamical interaction between recurrent microcircuits, which could have a key role in inference, memory and behavior.

1. Introduction

An obvious aspect of information processing in the brain is the fact that every sensorial input is assimilated over an already existing internal "mental state" (Scholl, 2001). There is also evidence that most of the neural microcircuits are recurrent and that connectivity between cortical neurons is really high (on the order of $10^3 - 10^4$) (Stevens, 1989). Moreover, it has been shown that highly irregular spike trains can be obtained with fixed random connectivity networks, questioning the source of the neural noise (Brunel and Hakim, 1999; Fusi et al., 2000; Gerstner and Kistler, 2002). Within this context, we might attribute, at least partially, the apparently noisy, spontaneous neural activity to the "mental state" of a recurrent microcircuit (the so-called network effect).

It seems essential to analyze the neural activity of real neurons from a new computational perspective, which can reveal the dynamical nature of processing. Such an approach is really hard, since we deal with enormously large dynamical systems. Nonetheless, extensive research is already emergent, a good example being the work of Maass and Markram (Maass and Markram, 2003). Within this context, an exact link between the function of an individual microcircuit neuron and its activity cannot always be found with traditional methods, its dynamics and interactions being far too complex for a comprehensive analysis. An alternative approach is to look at the activity of a population of neurons rather than to single units. This in turn, has its

own limitations since one has to determine the target population (not always an easy task) and apply statistical analysis which might sometimes reveal irrelevant functional aspects.

An emergent new technique of studying such large, complex, dynamical neural systems is to build and test large scale models, with biologically plausible components (Maass, 2002a; Natschläger, 2002). Using this technique, a novel approach to understanding the activity of high dimensional recurrent microcircuits has been used in the "liquid state machine" (LSM) (Maass, 2002b) and "the echo state approach" (Jaeger, 2001). In both cases, a population of readout neurons is trained to observe the highly dynamic processing in a recurrent module and convert the unstable dynamics into stable, reliable outputs.

Although, the recurrent microcircuit is highly biologically plausible, the existence of separate, independent readout populations (Maass 2002a) in the brain, that can be trained by regression methods, seems highly unplausible. More likely, there is a continuum of microcircuits interconnected in a complex, evolution guided manner, where parts of a recurrent microcircuit are readout for other microcircuit modules and vice-versa. Plausible mechanisms like STDP could be the basis for the interoperability of different spatial modules of the continuous microcircuit sheet. Development could be guided by behavioral experience, by the interaction between the environment and the organism (Steels & Brooks, 1995; Pfeifer & Scheier, 1999; Florian, 2003), in addition to genetical constraints. Dynamic associations might arise while the brain could self-wire itself in a complex way, perhaps too complex for a comprehensive analysis. Within this framework, we questioned the existence of a strong, unifying principle that guides cortical processing, called "The Coherence Principle" (Mureşan, 2003).

2. Coherence and microcircuits

Coherence is a universal yet simple principle, which states that a given interpretation is only "sound" in a given context. Thus, we can never separate the interpretation from the context. Also, an important observation is the fact that neural activity is most of the time causal, i.e. the activity of each neuron depends on the context of afferent activity. Moreover, we have to take into account the increasing complexity of neural activity, as one proceeds upstream in a sensorial processing hierarchy in the brain. While the primary layers seem to have deterministic responses, mainly driven by the input (as is the case in the primary visual cortex), deeper processing structures have a complex activity with complex firing patterns of the neurons. Neuroscientists know very well that as one proceeds to higher levels of sensorial processing, the complexity of interactions between neural areas grows significantly. We could then separate the afferent populations of a cell into "input populations" and "context populations". Clearly, this separation depends on the "interpretation" that is expected from the cell and there are many possible ways for segregating these two classes. A simple and maybe plausible segregation would be to label as input populations all the afferent neurons that are placed "before" the cell in the sensorial processing chain. Then, context populations would be all the other neurons that contribute to the activity of the target neuron. Lateral interaction neurons belong, in this case, to the context class. This definition can be extended from single neuron targets to target populations.

The same input could trigger a whole different range of interpretations for different context activities. Thus, the activity of a cell or population that has important afferent activity, different from the input, cannot be assumed to be simply a function of the input. This is why a transfer function to characterize the activity in

higher visual areas, such as V4 or IT, based on the input stimulus, is very hard if not impossible to define.

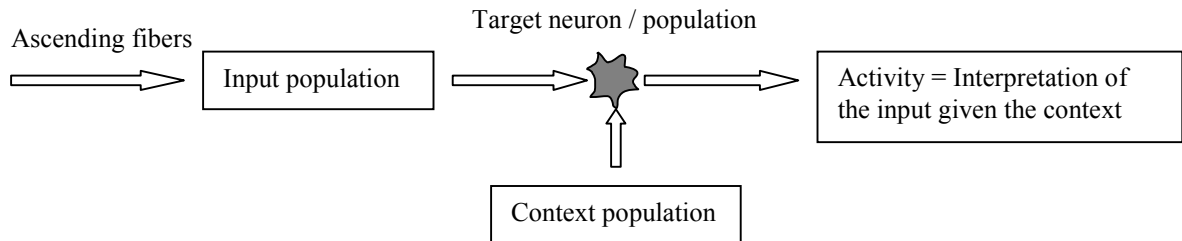


Fig. 1. Input / context segregation from a neuron or population's perspective

Microcircuit populations might cooperate to encode complex contexts and modulate other populations to generate interpretations. Competition, which has been extensively used by modelers, might not be a real plausible mechanism. Instead, we could have the following case: many populations cooperate, one being the observer and the others its input (simplified case). The observer population could learn to interpret only the input that is coherent with its current context, disregarding the other inputs. We could say that the relevant input wins the observer's attention, although no explicit competition takes place between the input populations (like lateral inhibition, etc).

Clearly, the principles presented here are very general but very abstract too. Many questions have to be answered before validating or perhaps invalidating them:

- how can microcircuits cooperate?
- which is the relevance of the recurrent dynamic processing within a microcircuit?
- are biological mechanisms fit for such kinds of processing?
- how do learning and self-organization emerge?
- how can such systems be analyzed? are analytical or statistical methods fit for such analyses?

Answering these questions is clearly not an easy task. However, a good starting point would be to use computer simulations of biologically plausible neural systems.

The analysis, largely empirical, starts with simple models and extracts various aspects from these simulations. Conclusions gathered in these steps could help neuroscientists to design experiments and also interpret neural data in a different, non-conventional manner.

3. A microcircuit "at work"

In this section we will focus on analyzing some aspects of processing with perturbations on a generic recurrent microcircuit. Although we have stated that isolated, specialized readout populations are highly implausible, a good way to start is to use such a simplification. At the same time, we are trying to build highly plausible models that have a value both in experimental and computational neuroscience.

Neural models with rich dynamics, close to their biological counterparts, but still computationally effective, have to be used. One such neural model, developed by Izhikevich (Izhikevich, 2003), has a very efficient implementation and it is, at the same time, very rich in dynamics. At the synaptic level, Senn and Markram have developed a highly plausible algorithm for modifying the neurotransmitter release probability (Senn and Markram, 2000).

Based on such plausible building-blocks, we have constructed a neural architecture that consists of 4 distinct populations of neurons. The first-one, also called input, is made up from a single layer of neurons that receive external current drive, injected by the experimenter.

The input population projects onto the first layer of a four-layered recurrent microcircuit with random connectivity. We used a mix of chattering and resonator neurons, 20% being inhibitory and the rest excitatory (see the model of Izhikevich, 2003).

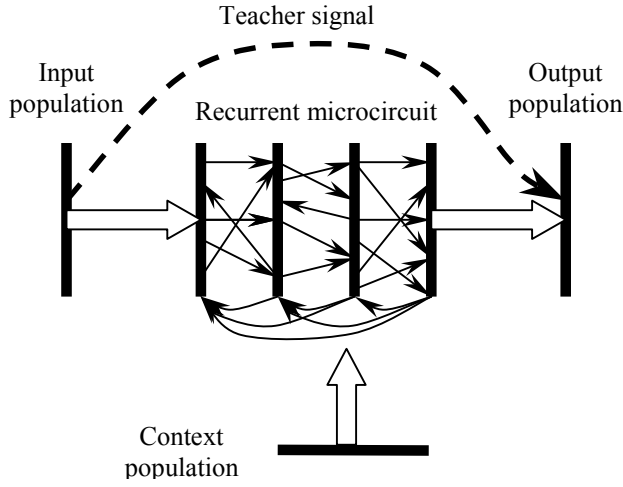


Fig. 2. The architecture of the test model.

The synapses are dynamic in the sense that the post-synaptic current is delivered using the following set of equations:

$$PSC(t) = A_{syn} \cdot W_{syn} \cdot g_{syn}(t-1) \cdot (E_{syn} - U_{post}(t-1)) \quad (1)$$

$$g(t) = g(t-1) \cdot e^{-1/\tau_{syn}} \quad (2)$$

where, A_{syn} is the absolute maximal synaptic strength, W_{syn} is the synaptic efficacy (between 0..1), g_{syn} is the time dependent conductance, E_{syn} is the reversal potential for the synapse class (-90 mV for inhibitory and 0 mV for excitatory synapses), U_{post} is the membrane potential of the post-synaptic neuron. Each time a pre-synaptic spike occurs, the value of g is incremented by 1, allowing for the superposition of synaptic stimulation effects.

Within the microcircuit, the synapses are randomly generated but still, retinotopy is maintained (a synapse from a neuron at a given location usually projects to a neuron in a different layer but in a given range of corresponding distance). Also, lateral interconnections between neurons in the same layer exist. Connectivity is

mainly inspired from the rules of connectivity determined for the visual system (Churchland and Sejnowski, 1999).

The output population is connected through STDP modeled synapses to each neuron in the recurrent microcircuit. We used for STDP the algorithm of Senn and Markram (Senn and Markram, 2000). Also, there are a number of synapses with the input population, that are activated only during the so-called "self-training", i.e. the input is injected also into the output to provide a teacher signal. The whole architecture is designed in order to simplify the experimental procedure and analysis, hence it might not be fully plausible. For example, the teacher signal might come from other microcircuits that have their own dynamics, but for the sake of simplicity, we provide an explicit teacher signal (the input in this case). Also, the output population might as well be part of another dynamic microcircuit or the same microcircuit (this is a clearly problematic separation).

An additional population, that in reality should be a microcircuit, is provided in order to inject a wave of spikes into the microcircuit. This population is called "the context population". The context population is controlled by injecting external currents during the experimental procedure. Neurons in the context population are connected via weak inhibitory synapses to the microcircuit population neurons, in a random manner.

3.1. The dynamics of the model

The membrane oscillations of the resonator neurons, the chattering neurons and the highly recurrent synaptic connections, cause the microcircuit to have a self-sustained activity, even in the absence of input. We call the activity of the microcircuit at any given time, the microcircuit's "mental state". Any spikes from the input or context populations generate a perturbation of the microcircuit's mental state.

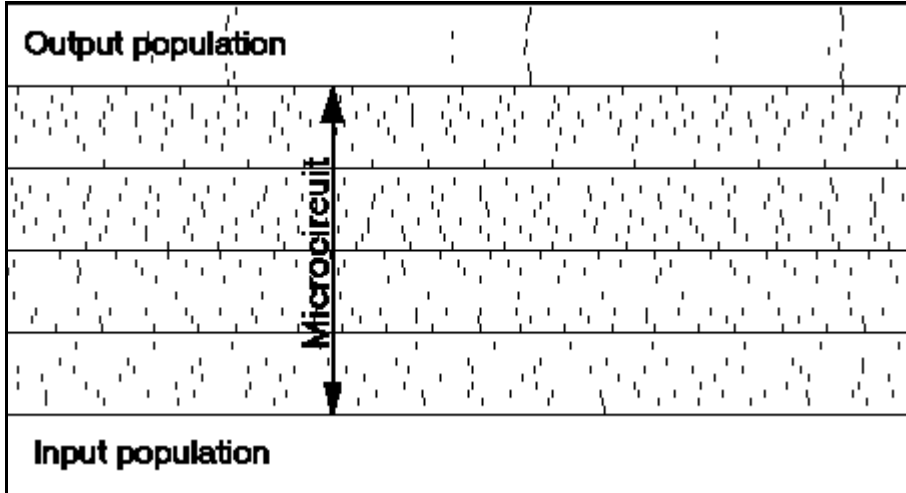


Fig. 3. Self-sustained spontaneous activity of the microcircuit in the absence of input.

An important observation is that the processing performed by the microcircuit is continuous, between two inputs, it has a sustained continuous "mental activity". Unlike traditional neural network models, where the input guides the processing (which does not take place in the absence of the input), the use of recurrent microcircuits opens a new, more plausible approach to neural processing.

Surprisingly, the spike trains generated by the recurrent microcircuit are highly variable, resembling a spontaneous random activity (although, no random currents are injected, like in other models).

3.2. Experimental setup

Even for this simple model, with only about 60 neurons, a good and safe analysis method is hard to find. Using rate analysis on the neurons inside the microcircuit is impractical and should be avoided since we are interested in the millisecond dynamics and each individual spiking. An empirical observation during experiments was that although the rate of a neuron did not significantly change, for very weak inputs, the output predictions of the population dramatically changed.

During experiments, a frequency-modulated signal is injected in the input layer. We used two types of signals, a "convex" signal and a "concave" signal (see Fig. 4).

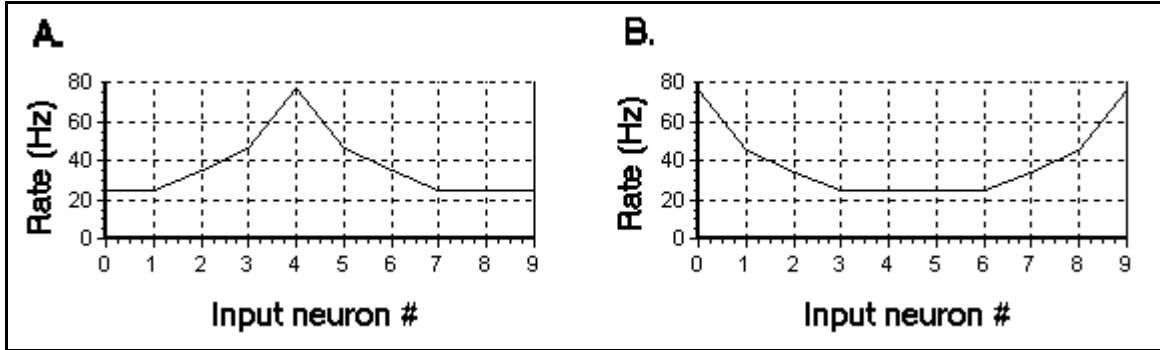


Fig. 4. The two types of input signals used in the experiments. A. "Convex" rate profile; B. "Concave" rate profile.

After presenting each input to the system, the activity pattern in the output population is analyzed. We empirically observed a consistent (and probably expected) transformation of the input rates into output phase coded responses. Consistent "convex" and "concave" activity profiles (encoded in phase) emerge in the output population as a result of the "convex" and "concave" input rates, respectively. The phase coded response can be observed in fig. 5.

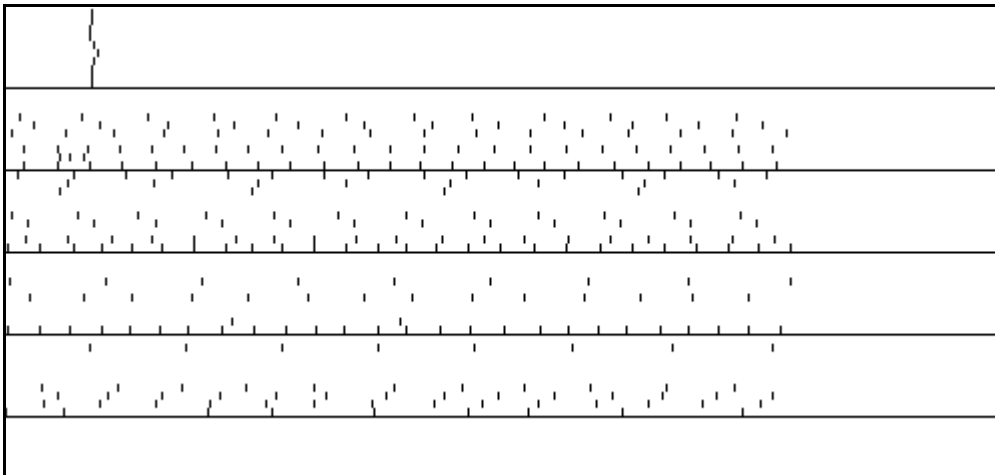


Fig. 5. An example of convex, phase coded activity of the output population, after shutting down the inputs.

Rate analysis reveals no frequency-coded responses in the output population, not even during the stimulation period. The phase-coded responses appear only after

shutting down the input. This means that the output population learns (through its STDP synapses) to interpret the microcircuit's spontaneous activity in a certain way (convex or concave).

As a measure of convexity of the output responses, we used the following formula:

1. during 500 ms, the number of "convex" / "concave" events are summed-up (for each convex event 20 was added to the convexity measure (CM), for each concave event, 20 was subtracted from CM);
2. an event was only validated if all the output neurons fired in a window of 20 ms length; if the relative phases of the firings respected an approximated convex profile, the event was considered to be "convex"; similarly, if the profile was approximated as concave, a "concave" event was recorded.

During the first experiment, a convex input has been presented for about 40 seconds and then, after resetting the simulation, a concave input has been presented for the same number of seconds.

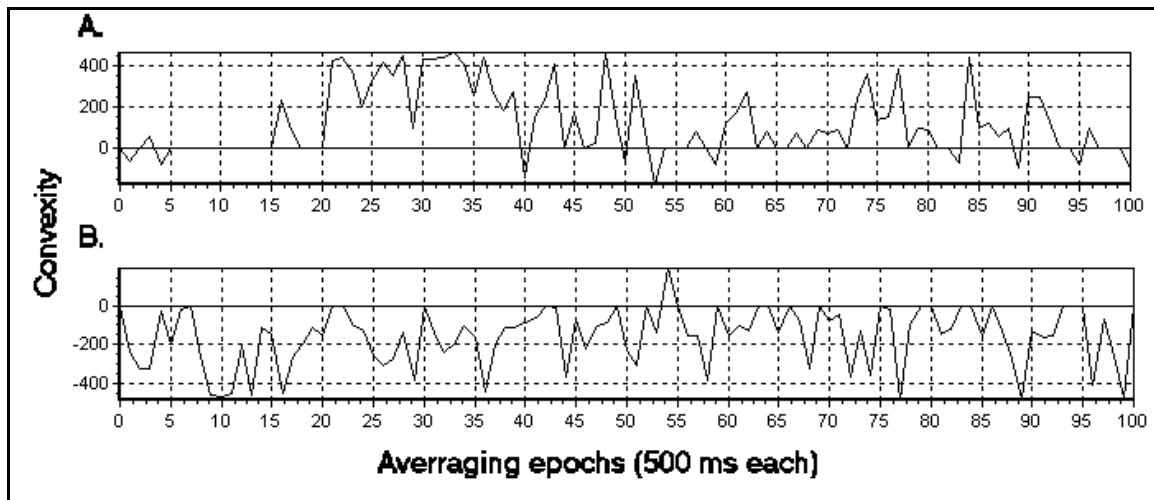


Fig. 6. Convexity measure for the "convex" task (A) and "concave" task (B). Recall events are highly irregular in time, but consistent (always, with few exceptions due to the measuring method, convex or concave).

Results, presented in fig. 6, clearly show that the output population interprets the reverberating activity (after stimulus shutdown) in a consistent way, proving that the convexity measure is good for analysis. Moreover, the high irregularity of recall events emphasizes the complexity of the neural activity and the richness of its dynamics (we should bare in mind that no random components were added to the model).

More intriguing results are obtained if the system is presented with both convex and concave inputs in two successive training epochs. The behavior of the system is in this case highly unpredictable, alternating, episodic recalls of "convex" or "concave" events occur with high irregularity (or it might be that the analysis method fails to reveal some regularity). If the second training epoch is prolonged, the system's recall will become deterministic, favoring the last stimulus. We should mention that in the absence of any past input (training epoch), the system also behaves randomly but eventually settles to a 0 convexity recall.

Finally, the most important experiment tries to reveal some possible mechanisms of context modulation and conditioning. The context population is activated (firing spike trains with a mean frequency of 20-50 Hz) during the "convex" training epoch. Then, it is shut down. After this step, the system is allowed to freely evolve for 50 seconds. A "concave" training epoch follows for 60 seconds. After the two training epochs, the activity of the output population is observed.

In the beginning, recall is mainly biased towards the last type of stimulus presented (concave) during the training epochs. When the context population is activated (firing random spikes with a mean frequency of 20-50 Hz), the associated stimulus is recalled (the convex stimulus). The context signal has no predefined structure (other than a quite stable frequency) thus the switch towards the recall of the associated stimulus (convex) is not due to some spurious interactions with the output population.

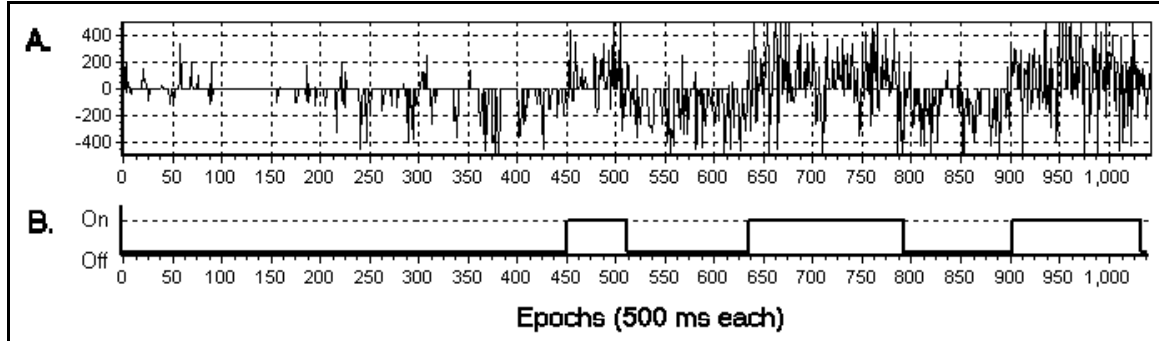


Fig. 7. The system's behavior during the context experiment. A. The convexity measure of the output activity. B. The activation state of the "context population".

The implications of these findings are important since they show how a simple population of neurons is able to store and preferentially retrieve dynamic information presented in the past by using an association context. The enormous complexity of interactions in the brain is another important aspect that is emphasized by the study of such models.

4. Conclusions

The present study tries to reveal some new ways of looking at neural processing with recurrent microcircuits. It seems essential to understand the complexity of neural interactions from a dynamical perspective, where each microcircuit module interacts with other microcircuits, yielding context modulation effects and the perturbation of the "mental states". Although the complexity of such systems is far beyond our limited analysis capabilities, the presence of a unifying principle (like coherence) could help researchers understand the way neural modules cooperate.

A good method of studying large dynamical systems is to use biologically plausible models, which could prove valuable, both from a modeling and experimental perspective. The behavior and dynamics of such models could reveal

the functional relevance of many biological principles of organization and molecular / cellular mechanisms.

From a computational perspective, we have to re-evaluate the source of neural noise and understand that at least in part, apparently hazardous neural dynamics, are in fact coherent "mental states" and emerge from the network dynamics.

The complexity and richness of the interactions between neural microcircuits can only be understood from a dynamical perspective. The experimental evidence and biological mechanisms all indicate that the ongoing processing in the brain is essentially temporal and dynamic, providing at the same time the basis for future analyses both in computational and experimental neuroscience.

Short and long-term memory might be based not only on synaptic plasticity but also on the dynamical interactions that are activated between multiple neural microcircuit modules (for example, the change at a single synapse, might trigger an entire excursion of the dynamical activity of a microcircuit which in turn contributes to encoding and memorizing the stimulus). Conditioning, might also be a micro-level phenomenon which relates to contexts and interpretations (coherence).

Acknowledgement:

I would like to thank to SC. NIVIS SRL (www.nivis.com) for supporting this research and to Prof. Leonardo Chelazzi from the University of Verona for introducing me the experimental neuroscience field.

References

Brunel, N. and Hakim, V. (1999). Fast global oscillations in networks of integrate and-fire neurons with low firing rates. *Neural Computation*, 11:1621-1671.

- Churchland, Patricia S., Sejnowski, T.J. (1999). *The computational brain*. MIT Press, Cambridge, MA.
- Florian, R.V. (2003). Autonomous artificial intelligent agents. *Technical Report Coneural-03-01*, (www.coneural.org).
- Fusi, S., Del Giudice, P., and Amit, D. J. (2000). Neurophysiology of a VLSI spiking neural network: Lann21. *In Proceedings of IJCNN 2000*.
- Gerstner, W. and Kistler, W. M. (2002). *Spiking neuron models*. Cambridge University Press, Cambridge, UK.
- Izhikevich, E.M. (2003). Simple model of spiking neurons, *IEEE Transactions on Neural Networks*, (in press).
- Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks. *GMD Report 148, German National Research Center for Information Technology*, 43 pp.
- Maass, W., Natschläger, T., Markram, H. (2002). A fresh look at real-time computation in generic recurrent neural circuits. Submitted for publication.
- Maass, W., Natschläger, T., Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531-2560.
- Maass, W., Natschläger, T., Markram, H. (2003). Computational models for generic cortical microcircuits. In: *Computational Neuroscience: A Comprehensive Approach*, J. Feng, Ed., CRC-Press.
- Mureşan, R.C. (2003). The Coherence Theory: Simple Attentional Modulation Effects. *Computational Neuroscience Meeting 2003 Proceedings*.
- Natschläger, T., Maass, W. (2002). Spiking neurons and the induction of finite state machines. *Theoretical Computer Science: Special Issue on Natural Computing*, 287:251-265.

- Pfeifer, R., Scheier, C. (1999). *Understanding intelligence*. MIT Press, Cambridge, MA.
- Scholl B.J., Leslie A.M. (2001). Minds, Modules, and Meta-Analysis. *Child Development*, Vol. 72, 3:696-701.
- Senn, W., Markram, H., Tsodyks, M. (2000). An algorithm for modifying neurotransmitter release probability based on pre-and postsynaptic spike timing. *Neural Computation*, 13:35–67.
- Steels, L., Brooks, R. eds. (1995). *The artificial life route to artificial intelligence: Building embodied, situated agents*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Stevens, C.F. (1989). How cortical interconnectedness varies with network size. *Neural Computation*, 1:473-479.