ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 231-236

An algebra for recognition of spatio-temporal forms

Gilles Vaucher

SUPÉLEC
Av. de la Boulaie - B.P. 28
35511 Cesson-Sévigné Cedex - France
Gilles.Vaucher@supelec.fr

Abstract. In the new model presented here, post-synaptic potentials are considered as punctual numerical elements on which neurons perform operations. Modelled in this manner, a neural network appears like a calculating machine with an asynchronous internal architecture. To formalize the functioning of such a calculator we propose an algebra of impulses based on the capacities in sequence recognition of a neuron (W. Rall's biological model). The aimed applicative domain is the realisation of perceptive systems which exploit signal dynamics.

1. Introduction

The realisation of man-machine interfaces based on natural means of communication utilised by human being pose the problem of mastering the treatment of spatio-temporal forms. When we represent by a series of vectors, composed of numerical values, a sequence made of audio signal spectra or a sequence of images of a visual scene it is possible to study their dynamics by treating the successive variations of vectors (delta coding). These series of vectors made of positive or negative pulses are spatio-temporal forms which we identify using the metric properties of an appropriate vector space.

2. Sequence recognition at neuron level

The followed approach is *neuromimetical.* It is based on a simplified biological model of dendritic tree which Wilfrid Rall proposed in 1962 [4]. According to [2], this *ball and stick* model is still a good first approximation of passive electrical properties of certain neurons. With this model, Rall studied the delays and deformations to which the post-synaptic potentials (PSPs) are subjected during their propagation towards the soma, as a function of their place of injection. He also studied the summation of PSPs and showed the influence of the combined effects of the position and the temporal order of the injections on the observed potential at the soma; by means of the addition of a threshold,

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 231-236

the neuron has the property of recognizing the sequences of impulses which are close to a privileged sequence.

For benefiting from this property of the Rall model, while conserving the vectorial representation of a common artificial neuron, we propose in [8] a plane coding of elementary PSPs. We consider only PSP maxima (PSPM) and retain their date and amplitude (fig. 1.a). Then, in a plane representation, we associate to each PSPM a punctual PSP (PPSP) having for length and phase the amplitude and date of the PSPM (fig. 1.b). PPSP summation is defined as a sum of vectors; the propagation of a PPSP is represented by a continuous decrease of its amplitude and a rotation in the plane (fig. 2). We represent the sequences of activities arriving at neuron inputs by vectors; each component of these vectors has a plane representation with the current instant of time as the common time reference (fig. 3).

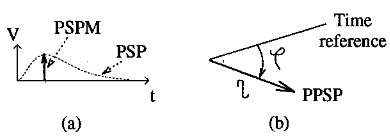


Figure 1: (a) Elementary PSP and PSPM. (b) Plane coding of a PPSP (η corresponds to the PSPM amplitude and φ to its phase).

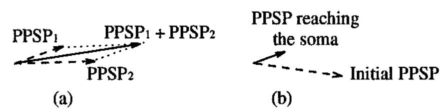


Figure 2: PPSP summation (a) and propagation (b).

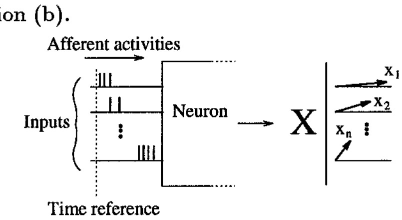


Figure 3: Vectorial coding of a sequence.

In the spatio-temporal artificial neuron (STAN) so modeled, X vector synthesizes, by temporal summation, the afferent activities synapse by synapse of a neuron. It is similar to the input vector of a common artificial neuron. It is the STAN short-term memory.

3. Formalizing with the help of complex numbers

3.1. The selected scalars

Vector X component formalizing needs scalars having two degrees of freedom. We, at first [6, 3], used the ordinary complex field ($\mathbb{C}$). But, by an analysis of the geometric properties of the plane, in particular those which are related to the choice of distance and angle measurements, we show in [7] that hyperbolic complex numbers set, noted here $\mathbb{C}_e$, is better adapted than $\mathbb{C}$ for representing PPSP. These numbers, less used than ordinary complexes, have a ring structure [10]. They are of two different types and have the form :

$$z = \xi\eta \exp e\varphi = \xi\eta(\cosh\varphi + e\sinh\varphi)$$

where $e^2 = 1$ (*one* and not *minus one*), $(\eta,\varphi) \in \mathbb{R}^2$, $\xi = 1$ for numbers of first type and $\xi = e$ for numbers of second type. The conjugate is defined as :

$$\overline{z} = \overline{a + eb} = a - eb \quad ; \quad (a, b) \in \mathbb{R}^2$$

3.2. Vector representation

We utilise hyperbolic complexes for coding the components of X :

$$X = \left| \begin{array}{c} \vdots \\ x_j \\ \vdots \end{array} \right. \quad \text{with} \quad \left\{ \begin{array}{lcl} x_j & = & \xi_j \eta_{x_j} \exp e\varphi_{x_j} \\ \varphi_{x_j} & = & \mu_T t_{x_j} \end{array} \right.$$

The value of ξ_j is 1 for excitatory inputs and e for inhibitory inputs; the amplitude η_{x_j} is positive or zero; t_{x_j} correspond to the delay which separate the date in the past associated to x_j and the current instant of time; μ_T is the inverse of a positive time constant.

Vector W, which corresponds to the weight vector in a common artificial neuron, is also represented in $\mathbb{C}_e^n$, n being the number of inputs. Its j^{th} component code the characteristics of the synapse j (inhibitory or excitatory type, efficiency and position). The type is the same as the one of the j^{th} component of X. Since W is of the same space as X, hence it codes a sequence of the same nature as X.

3.3. STAN potential

The potential v of STAN is defined by the aid of an hermitian product :

$$v = X^T \overline{W} = \sum_j v_j \quad \text{where} \quad v_j = x_j \overline{w_j} = \xi_j \overline{\xi_j} \eta_{x_j} \eta_{w_j} \exp e(\varphi_{x_j} - \varphi_{w_j})$$

It represents the spatial summation of v_j knowing that each v_j formalize the PPSP resulting from temporal summation of the pulses received at the input j.

Because the product $\xi_j \overline{\xi_j}$ is equal to $+1$ or -1 according to the j input type, the effect of excitation or inhibition on the potential results from a characteristic of scalars. The coding in $\mathbb{C}_e$, therefore, makes the use of a positive metric compatible with the introduction of inhibitory effects, which is an advantage over the common artificial neuron.

η_{w_j} and φ_{w_j} code respectively the efficiency of the synapse j and a delay which corresponds to the time needed by a PPSP to go from input j to STAN *soma*.

3.4. Summation and propagation operations

PPSP summation operation is formalized by a sum in $\mathbb{C}_e$. To represent the attenuation and the displacement towards the soma which each PPSP is sub-

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 231-236

jected to during time (t), we define the propagation operation in $\mathbb{C}_e$ by :

$$\mathcal{P}(v_j) = v_j exp(-\mu_S + e\mu_T)t \quad ; \quad \mu_T < \mu_S$$

The attenuation is exponential with a negative constant $-1/\mu_S$; the temporal rotation comes from the term $\exp(e\mu_T t)$. The propagation operation $\mathcal{P}$ induces in STAN an autonomous dynamic such as defined in [1] :

$$\mathcal{P}_{t_1+t_2} = \mathcal{P}_{t_1} \circ \mathcal{P}_{t_2}$$

The linearity of $\mathcal{P}$ as a function of time considerably simplifies the transposition of STAN dynamics into equations, as compared to a coding in $\mathbb{C}$.

3.5. Metric properties

Potential v represents the PPSP which synthesizes at a given instant all the activities from a recent past received by STAN. In the framework of our formulation in $\mathbb{C}_e$, we may predict with the aid of the propagation operation the value v_s which v will take when the PPSP will reach the soma :

$$v_s = \mathcal{P}_{-\varphi_v/\mu_T}(v)$$

v_s, which is a function of X and W, possesses interesting metric properties concerning the temporal characteristics of sequences. In fact, vector W, which synthesize synapse characteristics, may be considered as a result of a certain sequence of activities. If we construct the set $\mathcal{E}$ of the sequence X which are not different from W other than by the dates at which the activities arrive, we prove [7] that in $\mathcal{E}$ v_s presents a local optimum for $X = W$. This means that, in $\mathcal{E}$, the excitatory or inhibitory effect of an activity on v_s is maximum when it arrives, relatively to the others, at the moment for which it is expected in the sequence W.

Hence W synthesizes the synapse characteristics by expressing them under the form of a sequence privileged by the STAN.

3.6. Production of an output

Because v corresponds to a PPSP which has the time required for reaching soma as a temporal reference, STAN produces an impulse as soon as the phase of v becomes zero (under the effect of time passage). v hence becomes v_s. For avoiding unexpected outputs arising from the reception of sequences much different from W, it is necessary to add a threshold. We therefore find, at the formal level, the sequence recognition property of Rall's model.

To be able to put into equations groups of neurons, we have selected a common output function with a linear range (STAN normal behaviour) enclosed between a low non-linearity (threshold) and a high non-linearity (saturation). Lastly, we consider that the transmissions are without attenuation or delay; this last effect has been considered redundant with the mechanism in place at each input level.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 231-236

4. Putting into equations a group of neurons

Our aim is to use the algebra proposed above to put into equations groups of neurons for realizing given functions. To illustrate the approach we take as an example an architecture with lateral inhibitions to discriminate elementary movements in a sequence of images. We present the case of a mono-resolution of speed and consider four neighbour pixels that we project on a cluster of four STAN which inhibit each other mutually. Every time a pixel goes from white to black (binary image), a pulse is emitted on the four STAN corresponding input. To be able to detect elementary movements towards the left, right, up or down, discrimination must be made between four prototype sequences (fig. 4). For that it is sufficient to :

- determine vector W_k^E made of the excitatory *weights* of neuron k by imposing on it to be colinear at the prototype sequence χ_k^E,
- calculate the inhibitory *weights* by putting into equations the orthogonality (at near thresholds) of the prototype sequences in the output space of the cluster,
- choose thresholds in a manner such that unexpected outputs are avoided (these outputs may occur at the beginning of a sequence before the concerned STAN produces its output and inhibits its neighbours).

By proceeding in this manner, we thus obtain by simulation a perfect discrimination between the different types of movement when the speed of the moving object is nominal. This discrimination holds for a speed lying between half and double of the nominal speed.

Such an elementary group of neurons may be integrated in a multi-layer network. A second layer, for example, may be used for detecting coincidences or sequences of local movements. Each layer is put into equations by following the same proceeding.

5. Conclusion

Starting from Rall's dendritic tree model, we propose a STAN which is based on an algebraic approach. In STAN, the leaky integrator mechanism [5] and the concept of time delay at each input [9] are linked together within a formal framework. This artificial neuron possesses a short-term memory which varies continuously and also has the intrinsic capacity of identifying the end of a known sequence. The distance between sequences derives from metric properties. Thanks to the algebraic approach it is possible to put into equations groups of neurons which communicate asynchronously by impulse exchanges.

The prospective applications are perceptive systems which exploit signal dynamics (treatment of movement in images, handwriting recognition, speech recognition). We are currently working on such projects. We also aim to integrate this condensed temporal neuron model into a specific hardware.

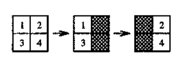
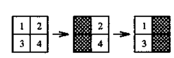
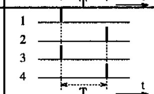
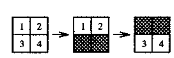
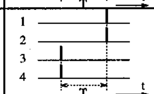
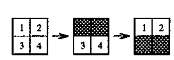
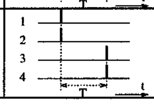
Move to	Four pixels	Pulse order	Prototype sequences
left			$\chi_g^E = \begin{pmatrix} I \\ \mathcal{P}_T(I) \\ I \\ \mathcal{P}_T(I) \end{pmatrix}$
right			$\chi_d^E = \begin{pmatrix} \mathcal{P}_T(I) \\ I \\ \mathcal{P}_T(I) \\ I \end{pmatrix}$
up			$\chi_h^E = \begin{pmatrix} I \\ I \\ \mathcal{P}_T(I) \\ \mathcal{P}_T(I) \end{pmatrix}$
down			$\chi_b^E = \begin{pmatrix} \mathcal{P}_T(I) \\ \mathcal{P}_T(I) \\ I \\ I \end{pmatrix}$

Figure 4: Corresponding to each elementary movement, we associate a prototype sequence. I is the amplitude of each input pulse. T is the delay which is between the beginning and the end of a sequence. It is reciprocally proportional to the object speed in the image. To compute the component values of a prototype sequence, the date at which the last pulse is received is used as the sequence time reference.

Acknowledgments

We wish to thank B. Picinbono and R. Baduel for giving us the opportunity to work on this project, J.C. Fort for his worthy comments and A. R. Baig for his assistance in translating this paper.

References

[1] P. Bourret, J. Reggia, and M. Samuelides. *Réseaux neuronaux*. Teknea, 1991.

[2] D. Johnston and S. M.-S. Wu. *Foundations of Cellular Neurophysiology*. MIT Press, 1995.

[3] N. Mozayyani, V. Alanou, J.-F. Dreyfus, and G. Vaucher. A spatio-temporal data-coding applied to kohonen maps. In *International Conference on Artificial Neural Networks*, volume 2, pages 75–79. EC2 et Compagnie, Oct. 1995.

[4] W. Rall. Core conductor theory and cable properties. In S. R. Geiger, E. R. Kandel, J. M. Brookhart, and V. B. Mountcastle, editors, *Handbook of Physiology : The Nervous System*, volume 1, chapter 3, pages 39–97. American Physiological Society, 1977.

[5] M. Reiss and J. Taylor. Storing Temporal Sequences. *Neural Networks*, 4:773–787, 1991.

[6] G. Vaucher. Un modèle de neurone artificiel conçu pour l'apprentissage non-supervisé de séquences d'événements asynchrones. *Valgo, ISSN 1243-4825*, 93-1:66–107, mai 1993.

[7] G. Vaucher. *A la recherche d'une algèbre neuronale spatio-temporelle*. PhD thesis, Supélec, déc. 1996.

[8] G. Vaucher. Neuro-biogical bases for spatio-temporal data coding in artificial neural networks. In Springer, editor, *Artificial Neural Networks — ICANN 96*, pages 703–8, July 1996. Lecture Notes in Computer Science 1112.

[9] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. Lang. Phoneme recognition using time-delay neural networks. Technical Report TR-1-006, ATR Interpreting Telegraphy Research Laboratories, 1987.

[10] I. M. Yaglom. *Les nombres complexes*. Dunod, 1966.