



HAL
open science

Les groupes polychrones pour capturer l'aspect spatio-temporel de la mémorisation

Regis Martinez, Helene Paugam-Moisy

► To cite this version:

Regis Martinez, Helene Paugam-Moisy. Les groupes polychrones pour capturer l'aspect spatio-temporel de la mémorisation. Deuxième conférence française de Neurosciences Computationnelles, "Neurocomp08", Oct 2008, Marseille, France. <hal-00331613>

HAL Id: hal-00331613

<https://hal.science/hal-00331613v1>

Submitted on 17 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

LES GROUPES POLYCHRONES POUR CAPTURER L'ASPECT SPATIO-TEMPOREL DE LA MÉMORISATION

Régis Martinez et Hélène Paugam-Moisy
LIRIS - UMR CNRS 5205
Université Lyon 2
5 avenue Pierre Mendès-France
F-69676 Bron cedex - FRANCE
email: {regis.martinez ; hpaugam}@liris.cnrs.fr

ABSTRACT

Nous nous basons sur le principe de la polychronisation, introduit par Izhikevich (2006), pour chercher une représentation synthétique de la trace générée en mémoire par des stimuli qu'un réseau de spiking neurones a appris à distinguer. Les spikes raster plots véhiculent une information trop dense. Nous extrayons de l'ensemble des neurones actifs ceux qui déclenchent des groupes polychrones. Nous retenons, comme représentation de la trace mémoire d'un stimulus, les groupes polychrones majoritairement activés pour la classe à laquelle appartient le stimulus. Nous montrons que cette représentation est suffisante pour distinguer les classes. Cette démarche rejoint celle du rank order coding proposée par Thorpe.

KEY WORDS

Spiking Neuron Networks - STDP - Polychronization - Spatio-temporal cell assemblies - Memory traces.

1 Introduction

Les progrès récents de l'imagerie médicale ont décuplé les recherches qui consistent à étudier comment le cerveau traite et mémorise l'information, et à chercher quelles régions cérébrales sont activées sous l'influence d'un stimulus. En neurosciences, on cherche à comprendre, par des expériences biologiques et des enregistrements, comment l'information est représentée et conservée en mémoire.

Les simulations de réseaux de spiking neurones de grande taille permettent aussi d'aborder une démarche d'explication des mécanismes de codage de l'information en mémoire. Le concept d'assemblée de neurones a été introduit par Hebb [1], au sens d'un groupe de neurones reliés par de fortes interconnexions excitatrices. L'idée est ensuite venue de considérer ces assemblées de neurones spatiales comme des sortes de "groupes de neurones grand-mère", chacune étant représentative d'un stimulus, voire même d'une notion ou d'un concept abstrait. L'intérêt des chercheurs s'est ensuite porté sur la notion de synchronisation des instants d'émission des spikes dans des sous-groupes de neurones, avec des travaux sur la

synfire chain [2] ou la *transient synchrony* [3]. Mais c'est seulement en 2006 qu'Izhikevich [4] dénonce le fait que la notion de synchronie reste trop restrictive pour exprimer la pleine puissance du traitement de l'information par assemblées de neurones. Tenant compte de la variabilité des délais de transmission de l'information entre neurones, il propose la notion de **polychronisation** qui, basée sur l'aspect temporel du traitement de l'information, permet d'expliquer une capacité de mémorisation quasi-infinie. Un groupe polychrone est un sous-ensemble de neurones dont les connexions ont des délais tels que l'activation, selon un motif temporel donné, de quelques neurones déclencheurs, provoque l'activation en chaîne de tous les autres neurones du groupe (cf. figure 1).

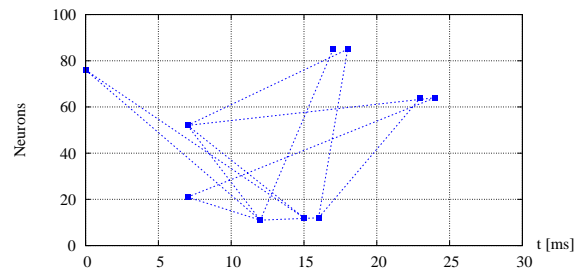


Figure 1. Exemple de groupe polychrone à 3 neurones déclencheurs.

Si l'on adopte un point de vue combinatoire, on peut étudier, selon différentes théories, la capacité de représentation d'un ensemble de n neurones, émettant au plus un spike chacun en réponse à un stimulus, dans une fenêtre de T millisecondes. Si le code consiste à compter les spikes émis, la quantité d'information maximale est de $\log_2(n+1)$. Pour un code binaire, elle serait de n . Si l'on adopte l'hypothèse du *rank order coding* [5], on peut monter jusqu'à $\log_2(n!)$, soit une quantité d'information de l'ordre de $n \log n$.

Plus généralement, si l'on inventorie, dans un réseau de N neurones, le nombre théorique d'assemblées de neurones spatiales possibles (i.e. le nombre de sous-

ensembles de neurones distincts), on obtient une capacité de $N!$. Mais si l'on adopte le point de vue d'Izhikevich, en considérant qu'un même neurone peut appartenir à plusieurs groupes polychrones simultanément, et si l'on considère des assemblées de neurones temporelles, alors la capacité devient théoriquement infinie.

Il est très séduisant de savoir expliquer que la capacité de stockage de l'information par le cerveau puisse être gigantesque, mais en contre-partie, cette information est assez déroutante. En effet, elle rend plus complexe que jamais la tâche qui consiste à comprendre comment sont stockées les informations en mémoire et comment deux stimuli distincts peuvent être différenciés. Le but de cet article est de s'appuyer sur un modèle artificiel de réseau de spiking neurones pour présenter une méthode permettant d'extraire une représentation synthétique et discriminative de l'activité neuronale provoquée par un stimulus. C'est une piste vers l'explication de la théorie de la mémoire à traces multiples [6].

2 Le modèle

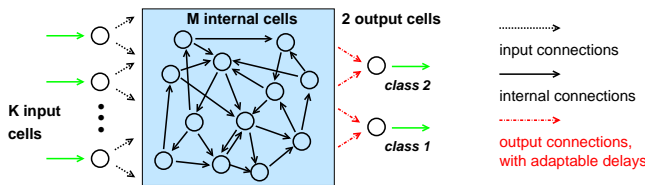


Figure 2. Modèle de réseau : on s'intéresse à l'activité neuronale à l'intérieur du réservoir (le carré bleuté).

Le modèle que nous avons développé [7] s'insère dans la famille des modèles de *Reservoir Computing*. Il s'agit (figure 2) d'un ensemble de neurones (le "réservoir") à connexions aléatoires et clairsemées, connecté d'une part à une couche d'entrée permettant d'injecter le stimulus et d'autre part à une couche de sortie (les *read-out* neurones) permettant de lire l'activité résultant de l'auto-organisation dynamique du réseau. Par l'ordre dans lequel ils émettent leurs spikes, les neurones de sortie indiquent la classe à laquelle appartient le stimulus. Dans le réservoir, les poids des connexions sont modifiés par plasticité synaptique et les délais des transmissions axonales sont aléatoires, mais fixes. Les délais des connexions qui vont du réservoir vers les neurones de sortie sont modifiés par apprentissage, selon une méthode originale, basée sur un critère à marge (temporelle), afin de différencier les instants d'émission des spikes des neurones de sortie : celui qui émet un spike en premier détermine la classe que le réseau donne en réponse au stimulus d'entrée. Pour résumer, d'un point de vue informatique, ce modèle est un classifieur neuronal à spikes dont la règle d'apprentissage se base sur deux échelles de temps : un apprentissage

local, à faible échelle de temps (1 ou 0,1 ms), par STPD, à l'intérieur du réservoir, et un apprentissage global, supervisé, à plus lente échelle de temps (de l'ordre de 100 ms), par modification des délais des neurones de sortie.

Tandis que dans l'article cité ci-dessus [7] nos préoccupations étaient plutôt de nature informatique, à savoir l'amélioration des performances du classifieur, celles que nous présentons dans cet article relèvent des neurosciences computationnelles puisque nous cherchons à comprendre les mécanismes qui sous-tendent les activités neuronales du réservoir. Nous cherchons à visualiser, dans un espace de représentation approprié, les traces mémoire qui permettent de distinguer les stimuli.

3 Représentation de l'activité neuronale

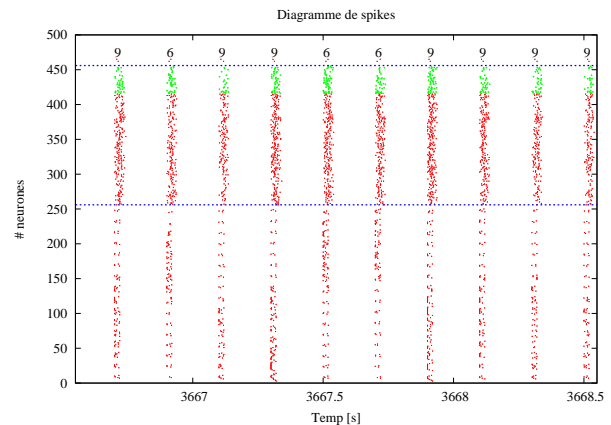


Figure 3. Exemples de *spike raster plots* pour la présentation (les 256 neurones du bas) de stimuli pour les chiffres 6 et 9 (classe précisée tout en haut), et activité neuronale dans le réservoir (neurones 257 à 417 excitateurs et 418 à 457 inhibiteurs) ; l'ordre d'émission des deux spikes de sortie (en haut) indique la réponse du réseau.

Le classifieur a permis d'apprendre à distinguer des chiffres manuscrits extraits des codes postaux américains, la base USPS¹ [8], avec des performances raisonnables : par exemple, 97% de succès pour discriminer des 6 et des 9. Ces chiffres sont des motifs visuels très réalistes : certains d'entre eux sont difficilement identifiables par un sujet humain et, pour chaque chiffre, la base de données présente une très grande variabilité. Lorsque l'on présente les chiffres après apprentissage, on constate qu'ils sont bien reconnus par le modèle, mais lorsque l'on regarde les activités neuronales à l'intérieur du réservoir, on a du mal à distinguer les raisons de cette bonne identification. Les *spikes raster plots* permettent d'affirmer que ce sont des assemblées de neurones temporelles, et non pas seulement

¹ <http://www-stat-class.stanford.edu/~tibs/ElemStatLearn/data.html>

spatiales qui semblent véhiculer l'information significative, comme nous l'affirmons déjà dans [9]. En effet, lorsqu'on a appris à distinguer un chiffre 6 d'un chiffre 9 (modèle à deux classes), on constate que la plupart des neurones sont activés à la fois pour les deux chiffres. Ce ne sont donc pas des sous-ensembles de neurones distincts qui les représentent en mémoire. Cependant les *spikes rasters plots* sont denses et présentent de fortes variations (cf. figure 3) : ils sont difficilement exploitables pour capturer les différences inter-classes et les similarités intra-classes. On pourrait peut-être tirer une information de ces rasters en leur appliquant des mesures de distance entre trains de spikes, telles que celles proposées par [10], mais nous avons poursuivi une autre piste.

Notre idée est de chercher à capturer les spikes les plus significatifs, un peu dans le même esprit que le calcul des *support vectors* dans le modèle SVM de Vapnik [11]. Dans cet esprit, nous pensons que les neurones qui pourraient jouer un tel rôle sont ceux qui (parmi les neurones activés lors de la présentation d'un stimulus) sont des neurones déclencheurs pour des groupes polychrones. En effet, leur activation engendre une bonne partie de l'activation des autres neurones observable sur les *rasters plots*, mais la connaissance des neurones déclencheurs est suffisante et elle procure une information plus synthétique.

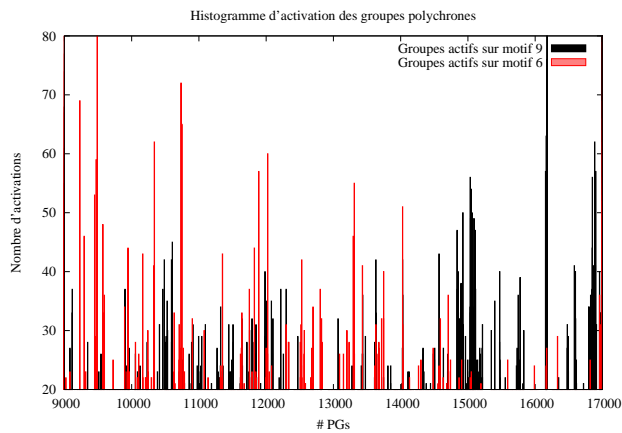


Figure 4. Histogrammes montrant, dans un intervalle représentatif (environ un tiers de tous les PG générés par la topologie du réservoir), les nombres de fois où les groupes polychrones sont activés par la présentation d'un 6 (rouge clair) ou d'un 9 (noir).

Nous avons donc effectué un changement d'espace de représentation en visualisant l'activité du réseau par les groupes polychrones activés, et non plus par les neurones activés. La figure 4 présente, sous forme d'histogramme, un sous-ensemble de tous les groupes polychrones possibles (plusieurs dizaines de milliers, même pour un réservoir de quelques centaines de neurones). En abscisse sont présentés les numéros des groupes polychrones (PG) ; un

trait vertical représente le nombre de fois où le groupe concerné a été activé lors de la présentation de tous les chiffres de la base USPS (de l'ordre de 600 motifs, pour chaque classe), en rouge s'il s'agit d'un 6 et en noir s'il s'agit d'un 9. La figure ne représente que les PG dont l'activité est reproduite au moins 20 fois sur tous les chiffres d'une classe donnée. On observe une spécialisation des PG activés, selon qu'il s'agit d'un 6 ou d'un 9, lorsqu'on présente les chiffres en test (STDP désactivée) après une phase d'apprentissage de 3645s de temps biologique simulé, soit 14 passes de la base d'exemples.

En partant de cette représentation dans l'espace des PG (les histogrammes de la figure 4), il est alors possible de classer les PG en plusieurs catégories, selon qu'ils ont ou non une classe préférentielle et selon la nature de cette classe. Nous affectons une classe à un PG (ici : 6 ou 9) si le groupe est activé au moins 10% de fois plus souvent pour cette classe que pour l'autre, sinon ce groupe est considéré comme étant dans la catégorie "sans préférence". Dans l'expérience relatée ici, nous avons obtenu 452 PG de catégorie 6 et 96 PG de catégorie 9, sur plus de 20000 PG générés par la topologie du réservoir.

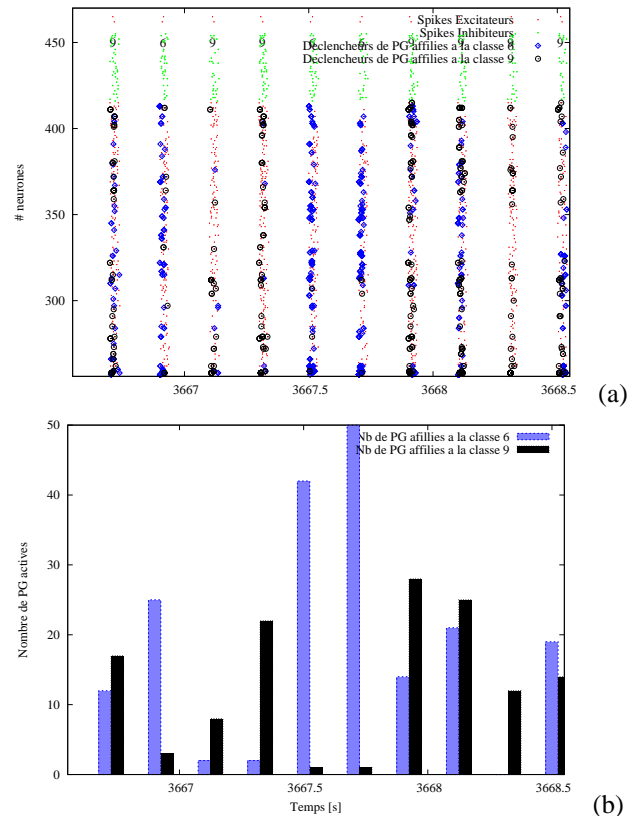


Figure 5. (a) *Spike raster plot* montrant les déclencheurs des groupes polychrones activés. On distingue les spikes déclencheurs des PG de catégorie 6 (losanges bleus) de ceux de catégorie 9 (ronds noirs). (b) Histogrammes des groupes polychrones activés, en correspondance avec les stimulations testées sur la figure 5a.

La figure 5a est une reproduction du spike raster plot précédent (figure 3), restreint aux neurones du réservoir et aux neurones de sortie, mais sur la même période de temps. Cette figure met en valeur les spikes qui sont des déclencheurs de PG, en distinguant les groupes de catégorie 6 (ceux dont la classe préférentielle est le 6 - losanges bleus) et les groupes de catégorie 9 (ceux dont la classe préférentielle est le 9 - ronds noirs). De nombreux spikes du raster sont donc sans intérêt (ceux qui restent affichés en rouge), et ce pour deux raisons : soit ils ne sont pas déclencheurs de PG, soit ils correspondent à des PG de la catégorie “sans préférence”. Il semble alors que l’activité du réseau puisse être représentée par une sous-activité synthétique, caractéristique d’une classe ou d’une autre, par le biais des déclencheurs de PG.

Pour plus de rigueur, nous quantifions cette activité par rapport à chaque classe : nous comptons, pour chaque stimulation, le nombre de PG affiliés à chacune des deux classes. Nous pouvons ainsi générer, pour chaque stimulation, un histogramme représentant la balance des groupes des deux affiliations différentes (figure 5b). L’histogramme nous permet de confirmer que, dans la réponse à une stimulation, la proportion de PG activés la plus importante est celle correspondant à la classe de la stimulation. Dans certains cas, qui restent rares (exemple de la stimulation présentée au temps 3668.5 s, sur la droite de la figure 5), la tendance est inversée. Ceci s’explique par l’observation de la réponse des neurones de sortie qui montre que ce motif est peu clairement discriminé (la marge temporelle n’est pas atteinte) par le réseau. Dans un souci de clarté, nous avons présenté les résultats pour seulement dix stimulations, mais les mesures réalisées sur d’autres intervalles de temps conduisent à des résultats analogues.

4 Discussion

Ce nouveau mode de représentation est une piste pour expliquer les mécanismes de traces spatio-temporelles des motifs en mémoire. Il permet de visualiser, de manière spatiale, mais dans l’espace des groupes polychrones et non plus dans celui des neurones, la différence des activités spatio-temporelles du réseau en réponse à des stimuli qu’il a appris à distinguer. Les groupes polychrones, dans leur définition-même, contiennent une information temporelle puisqu’ils ne sont activés que si leurs neurones déclencheurs reçoivent des spikes selon un motif spatio-temporel bien précis, défini par les délais de transmission des connexions afférentes. Par ailleurs, les activations des PG sont conditionnées par les poids des connexions (les poids excitateurs doivent renforcer suffisamment le signal entrant pour qu’il y ait émission d’un spike). Elles résultent donc de l’expérience acquise par apprentissage, mais ne seront réactivées selon une configuration semblable que si un stimulus similaire est présenté en entrée.

En ne s’intéressant qu’à certains spikes véhiculant une information significative, cette démarche rejoint celle du *rank order coding*, mais elle est néanmoins différente. En effet, nous prenons en compte non pas tous les premiers spikes émis, mais les spikes qui, parmi les premiers émis, sont des déclencheurs de groupes polychrones. Cette méthode nous paraît plus significative car l’activité neuronale contient toujours une part importante de bruit et de nombreux neurones, parmi ceux qui émettent en premier, peuvent très bien ne véhiculer aucune information significative. En revanche, la probabilité pour qu’un triplet (ou quadruplet, ou plus) de neurones déclencheurs d’un PG soient tous activés par simple effet de bruit est beaucoup plus faible. Ainsi la représentation que nous proposons nous semble plus pertinente pour expliquer, et surtout pour visualiser, les mécanismes de codage de l’information par assemblées de neurones temporelles, et, à plus long terme, pour étayer la théorie de la mémoire à traces multiples. Cette représentation rend compte de l’auto-organisation du réseau et, dans la plupart des cas, elle suffit pour identifier correctement la nature de la stimulation.

References

- [1] D.O. Hebb. *The Organization of Behaviour*. Wiley, New York, 1949.
- [2] M. Abeles. *Corticonics: Neural Circuits of the Cerebral Cortex*. Cambridge Press, 1991.
- [3] J.J. Hopfield, C.D. Brody, and S. Roweis. Computing with action potentials. In M.I. Jordans, M.J. Kearns, and S. Solla, editors, *NIPS*1997, Advances in Neural Information Processing Systems*, volume 10, pages 166–172. MIT Press, 1998.
- [4] E.M. Izhikevich. Polychronization: Computation with spikes. *Neural Computation*, 18(2):245–282, 2006.
- [5] S. Thorpe, A. Delorme, and R. Van Rullen. Spike-based strategies for rapid processing. *Neural Networks*, 14:715–725, 2001.
- [6] D.L. Hintzmann. “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, 93:411–428, 1986.
- [7] H. Paugam-Moisy, R. Martinez, and S. Bengio. Delay learning and polychronization for reservoir computing. *Neurocomputing*, 71(7-9):1143–1158, 2008.
- [8] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer-Verlag, 2001.
- [9] D. Meunier, R. Martinez, and H. Paugam-Moisy. Assemblées Temporelles dans les Réseaux de Neurones Impulsionnels. In *Actes de la conférence française NeuroComp 2006*, pages 187–190, Pont-à-Mousson, october 2006.
- [10] J.D. Victor and K.P. Purpura. Nature and precision of temporal coding in visual cortex: A metric-space analysis. *J. of Neurophysiology*, 76:1310–1326, 1996.
- [11] V.N. Vapnik. *Statistical learning theory*. Wiley, 1998.