

Résumé de la publication d'Anders Lansner, 2009

Associative memory models

from the cell-assembly theory to biophysically detailed cortex
simulations



Neurosciences computationnelles
Parcours SISN

Introduction

Anders Lansner est un professeur suédois à l'université de Stockholm. Il est officier dans le département de biologie computationnelle, ce qui le mène régulièrement à publier des articles sur ces études. En 2009, il publie "Les modèles de la mémoire associative : de la théorie de l'agencement cellulaire aux simulations biologiques du cortex"

Dans son article, il détaille l'évolution de la modélisation de la mémoire. De Hebb à la mémoire-attractrice, en passant par le modèle de Hopfield, il montre que cette zone du corps humain est un processus long et complexe à modéliser.

Modèles déjà existants

Pour bien comprendre cet article, il est nécessaire de comprendre le fonctionnement biologique de la mémoire. Cette dernière est traitée par le cortex associatif, qui correspond aux régions du cerveau chargées du traitement des informations. Dans la deuxième moitié du 20ème siècle, Donald Hebb pose les bases du concept de la mémoire associative corticale qui correspond en d'autres termes au processus d'assimilation et de stockage des informations. Ce phénomène est expliqué par un agencement particulier des neurones dans le cerveau et sa compréhension aujourd'hui encore n'est pas absolue.

Les recherches de D. Hebb ont ensuite permis à d'autres paradigmes plus poussés de voir le jour, toujours pour essayer d'expliquer plus en détail la fonction corticale. (comme le modèle de Hopfield en 1982 ou le concept de mémoire par attracteur qui est actuellement le modèle le plus étudié).

Le processus de la mémoire, et à plus petite échelle de la fonction corticale du cerveau est en réalité un phénomène holistique. Ça veut dire qu'on ne peut comprendre son fonctionnement qu'en regardant son intégration globale dans son milieu. Ainsi, une information qui circule entre les neurones crée des connections synaptiques. Mais alors, lorsqu'une information circule alors que les connections existent déjà ou partiellement, le stimulus envoyé au début va se refléter dans tout le réseau associé à l'idée originelle. Voilà donc le fonctionnement biologique de la mémoire. Ce phénomène explique aussi la reconnaissance des formes, qui est plus simple quand on a plus de détails, puisqu'alors, l'information va pouvoir passer par des chemins plus nombreux et trouver plus facilement l'information correspondant à la forme recherchée.

Modèle de Hebb

Les hypothèses de Hebb sont bien connues qui indiquent que la plasticité des synapses à long terme est un élément crucial à la formation de la mémoire.

La première proposition est que des neurones qui s'excitent ensemble se lient entre eux. Hebb pense que l'activation répétée d'une synapse en modifie la force, le "poids", et donc permet de faire circuler l'information plus vite d'un neurone à l'autre.

La deuxième proposition est que de tels processus contribuent à la formation d'assemblées cellulaires. Les groupes de cellules à plusieurs reprises co-activées deviennent liées

ensemble par des synapses excité : les synapses de Hebb (hebbian synapses). Ces objets mentaux supporterait la représentation de la mémoire et l'activité réverbérante et influencerait la dynamique de la perception ultérieure et l'appel à la mémoire.

La loi de Hebb s'écrit mathématiquement par la formule suivante :

$$w_{ij} = \frac{1}{P} \sum_{k=1}^P x_i^k x_j^k$$

Modèle de Hopfield

Le modèle mathématique de Hopfield est la manière la plus simple qui implémente le principe d'assemblée. Le réseau de Hopfield présente l'avantage de fonctionner avec des principes compatibles avec Hebb. De plus, il permet de constituer un modèle plus proche de la réalité biologique.

Un réseau de Hopfield inclus un ensemble de cellules connecté à tout. Lorsqu'il est utilisé en tant que réseau attractif de la mémoire, les connections du réseau sont entraînés par quelques règles d'apprentissage pour former la matrice de connexion.

Les équations communément utilisées pour mettre à jour la membrane d'activation et les valeurs de sortie de la cellule sont :

$$C_j \frac{\partial u_j}{\partial x} = \sum w_{ij} v_i - \frac{u_j}{R_j} + i_j$$

$$v_j = \tanh\left(\frac{u_j}{T}\right)$$

où u et v représente l'activation membranaire et v la valeur de sortie, respectivement, i et j indexent les unités pré et postsynaptiques. C et R représentent la capacité et la résistance de la membrane. Wij est un poids synaptique. En se donnant un ensemble de P patterns à stocker dans la mémoire avec N cellules, une règle d'apprentissage classique est la suivante :

$$w_{ij} = \begin{cases} 0 & \text{if } i = j \\ \frac{1}{N} \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu & \text{otherwise} \end{cases}$$

où epsilon k représente la sortie (+-1) de la cellule k en apprentissage de pattern mu.

Modèle de neurones à pic de fréquence

Dans les années 70, les scientifiques ont réalisé des réseaux de neurones basés sur le **modèle des neurones à pics de fréquence**. Ce modèle plus réaliste que celui de Hopfield permet d'étudier la fréquence des pics d'activité des neurones lors du passage de

l'information. Les tendances montraient une synchronisation des pics autour de 50Hz et il se trouve que cette synchronisation a été retrouvée dans le cerveau d'un chat quelques années plus tard. Ainsi, même si ce modèle ne fournit pas des résultats parfaits notamment sur l'organisation des cellules, et la réverbération de l'information lors du rappel d'un souvenir, il utilise quand même de nombreux fondements neuroscientifiques qui reflètent des comportements similaires au cerveau animal.

Lien entre modèles théorique et biologie

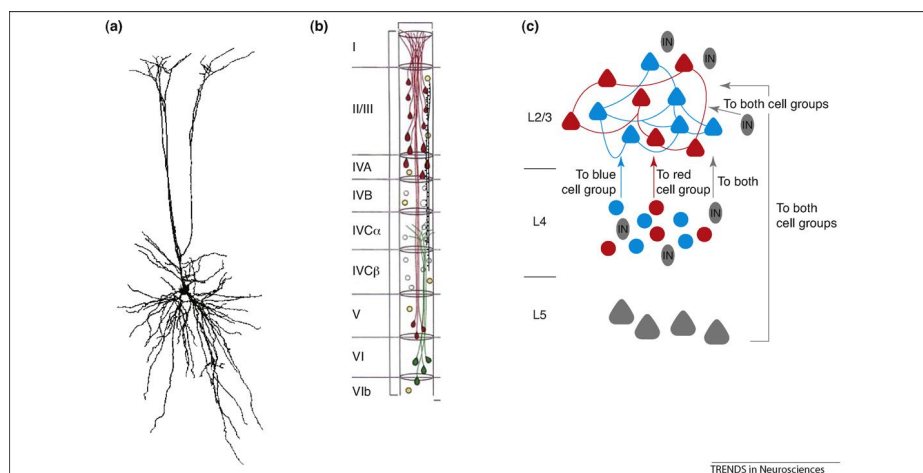
A quoi correspond biologiquement un réseau de neurones informatiques ?

Comme tout modèle, ceux décrits plus haut ne reflètent pas complètement la réalité. Par exemple Hopfield est assez éloigné de la neurologie. Il a notamment été critiqué sur plusieurs points :

- Une connexion multipoint à multipoint symétrique non fidèle à la biologie.
- il viole la loi de Dale, qui stipule que chaque neurone relâche seulement un neurotransmetteur à tous ses synapses.
- les niveaux d'activités sont trop élevés
- Le nombre de patterns possible à stocker est trop faible
- l'activité reste coincée dans les états attracteurs et les rappels avec des débits des neurones impulsionnels serait probablement plus lent qu'observer expérimentalement.

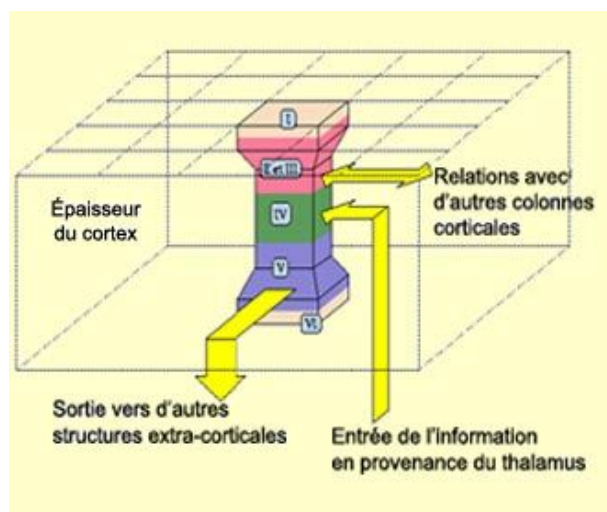
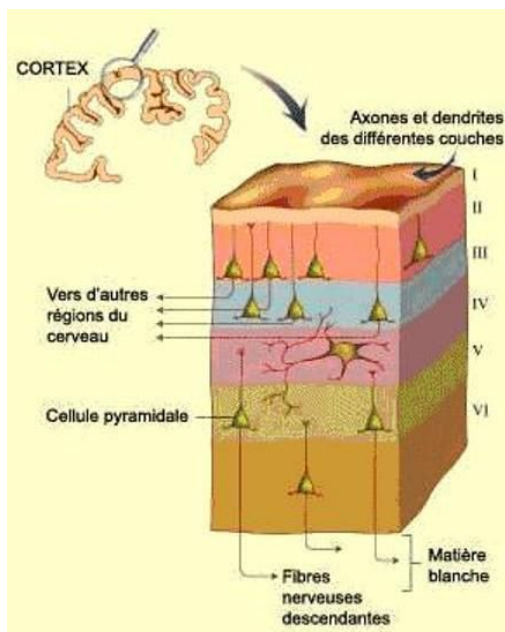
Ainsi on peut se demander ce que pourrait réellement représenter les cellules informatiques créées par ces réseaux de neurones virtuels ? 3 possibilités d'association entre les unités informatiques et la biologie sont possibles.

- un neurone simple, facile à visualiser conceptuellement
- une colonne du cortex cérébrale
- les couches 2-3-4-5 de cette colonne



En réalité, dans un réseau comme ceux décrits plus haut, une cellule informatique ne correspond pas à un neurone biologique. Le cerveau étant un système holistique, si le système ne représentait qu'un neurone, on n'aurait pas à rendre compte de l'intégration de cette cellule parmi les autres et le modèle échouerait à représenter le cerveau dans sa globalité. L'hypothèse la plus plausible est qu'une cellule corresponde au fonctionnement d'une petite population de neurone et plus particulièrement à une partie d'une mini colonne de la matière grise.

Cette colonne située dans le cortex, est un ensemble de neurones orientés verticalement en direction du cerveau. Leurs axones plongent plus ou moins profondément, ce qui provoque un effet de couche. Ces couches permettent ensuite de communiquer avec différentes autres parties du cerveau.



Mais le cerveau n'est pas encore suffisamment connu pour statuer clairement de l'impact de ces colonnes sur le réseau. Ainsi, on n'a pas encore de réponse à la question : "qu'est-ce qu'on modélise concrètement ?".

prévisions futures (informatique)

La puissance des machines et super-calculateurs qu'on a à notre disposition ne cessent d'augmenter d'années en années. Ainsi, d'ici quelques temps, les scientifiques pourront lancer de nouvelles expériences d'envergure pour simuler la mémoire et l'agencement des cellules. Si aujourd'hui les machines d'IBM réussissent à modéliser un réseau de la taille d'un cerveau de souris, les scientifiques s'accordent à dire que d'ici 15 ans, on sera capable de modéliser un cerveau humain. Ces prévisions s'appuient sur la loi de Moore qui prédit l'évolution des puissances de calcul des ordinateurs dans les années à venir. Cette loi est vérifiée depuis l'invention du microprocesseur et semble donc être plutôt fiable.

Pour info, le cerveau d'une souris compte 22-million de neurones connectés par 11-

Ainsi, nous serons en mesure d'étudier des architectures de neurones plus complexes. Les résultats de ces expériences feront alors le lien entre les observations biologiques (que nous avons aujourd'hui grâce aux méthodes d'études actuelles : EEG, MEG ...) et les résultats et prévisions théoriques.

Conclusion

Les recherches sur la modélisation de la mémoire avancent très rapidement aujourd'hui. Les modèles computationnels se perfectionnent au rythme des avancées informatiques ! Si les nombreuses simulations effectuées régulièrement permettent de faire avancer les modèles, les expériences et observations biologiques jouent également un rôle important dans le processus de modélisation pour ajuster les données. Cependant, la modélisation computationnelle du cerveau est sur une bonne voie, et le fait de pouvoir modéliser un cerveau humain d'ici 10 à 20 ans nous permettra à terme de comprendre les réactions localisées de notre cerveau, et donc prévenir de dysfonctionnements, ou autres maladies.