Plasticité des réseaux de neurones: Un modèle de Réseaux Booléens avec règle de Hebb

Jean-Paul Comet, Adrien Richard, Alexandre Muzy

Lab. I3S, Université de Nice-Sophia-Antipolis, France

5 novembre 2014

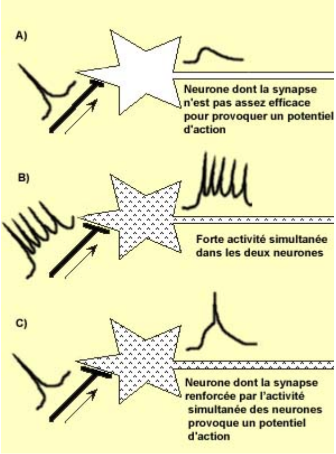
Introduction

- Règle ou théorie de Hebb (Donald Hebb, 1949).

  "Cells that fire together, wire together."
  "Des neurones qui stimulent en même temps,
  sont des neurones qui se lient ensemble"

- Mécanisme basique de plasticité synaptique : l'efficacité synaptique augmente lors d'une stimulation présynaptique répétée et persistante de la cellule postsynaptique.
- lorsque deux neurones sont excités conjointement, il se crée ou renforce un lien les unissant
- prix Nobel (2000) : Eric Kandel
  preuve de l'implication des mécanismes d'apprentissages de Hebb dans les synapses du gastropode marin Aplysia californica.

Long Term Potentiation : ...



1. Une exitation isolée du neurone post-synaptique ne suffit pas à déclancher la réponse
2. Une succession d'exitation permet de provoquer plusieurs fois le potentiel d'action
3. Après renforcement des liens, une exitation isolée suffit à déclancher la réponse

Un modèle de Hebb dans les réseaux booléens : Shih-Tsai

- $i, j \in [1 \dots n]$
- $x \in \mathbb{B}^n$
- $A \in \mathbb{R}^{n \times n}$
- $s_i^A = \sum_{j \text{ s.t. } x_j = 1} a_{ij}$ and $f_i^A(x) = \begin{cases} 1 & \text{if } s_i^A \geq b_j \\ 0 & \text{otherwise} \end{cases}$
- Les matrices évoluent au cours du temps : $(A^t)_{t \in \mathbb{N}}$
- Une **stratégie** : un choix de neurones à mettre à jours au temps t $I : \mathbb{N} \to \mathcal{P}([1 \dots n])$
  - **équitable** : $\forall i \in [1 \dots n]$, infinité de k tq $i \in I(k)$
  - **asynchrone** : $\forall n \in \mathbb{N}, |I(n)| = 1$

Les règles de Hebb (1)

- Système dynamique dans $(\mathbb{B}^n, \mathbb{R}^{n\times n})$
- Une trace : $(x^0, A^0) \to (x^1, A^1) \to (x^2, A^2) \to \dots$

$$\text{avec} \qquad x_i^{t+1} = \begin{cases} f_i^{A_t}(x^t) & \text{si} \quad i \in I(t) \\ x_i^t & \text{sinon} \end{cases}$$

- $\delta^t = x^{t+1} - x^t$
- $\Delta^t = A^{t+1} - A^t$

3 propriétés de la règle de Hebb vue par Shih-Tsai

I. Une trajectoire satisfait la règle de Hebb (I) sur un interval T si $\forall t \in T$

$$\begin{cases} \Delta_{ij}^t \geq 0 & \text{si} \qquad x_i^{t+1} = x_j^{t+1} = 1 \\ \Delta_{ij}^t \leq 0 & \text{sinon} \end{cases}$$

II. Une trajectoire satisfait la règle de Hebb (II) sur un interval T si $\forall t \in T$

$$|\Delta_{ij}^t| \geq |\Delta_{ji}^t| \text{ si } \exists t' \leq t \quad \text{tq} \quad (x_i^{t'} \neq x_i^{t''}) \wedge (x_j^{t'} = x_j^{t''} = 1) \quad \forall t' < t'' \leq t+1$$

si le dernier changement de i est antérieur à celui de j

III. Une trajectoire satisfait la règle de Hebb (III) sur un interval T si

$$\forall V \in \mathcal{P}([1 \dots n]), \sum_{i \in V} \min_{t \in T} a_{ii}^t \geq \sum_{i,j \in V} \max_{t \in T} |a_{ij}^t - a_{ji}^t|$$

l'auto-contribution est au moins aussi forte que l'asymétrie des contributions

Sur une trajectoire satisfaisant (I) qui boucle...

- soit $(x^0, A^0) \to (x^1, A^1) \to (x^2, A^2) \to \dots$
  une trajectoire satisfaisant (I)
- supposons qu'il existe $t_* < t < t^*$ tq $x^{t_*} = x^{t^*} \neq x^t$
- Alors

$$\sum_{t_* \leq t < t^*} \langle A^t x^t, \delta^t \rangle > 0$$

Somme des soma $\nearrow$ $>$ somme des soma $\searrow$

Preuve presque immédiate (les seuils ne changent pas)

Si on reste 2 pas de temps sur un même état...

1. soit $(x^0, A^0) \to (x^1, A^1) \to (x^2, A^2) \to \dots$
   une trajectoire satisfaisant (I)
2. supposons qu'il existe t tq $f_i^{A^t}(x) = x_i^t$ pour tout i
3. Alors

$$x^{t+k} = x^t \text{ pour tout } k \in \mathbb{N}$$

dès que la trajectoire bégaye, elle devient stationnaire

Preuve presque immédiate
(la modification de la matrice ne change pas la comparaison au seuil)

(asynchrone)+(I)+(II)+(III) ne permettent pas de boucler

- Soit $(x^0, A^0) \to (x^1, A^1) \to (x^2, A^2) \to \dots$
  une trajectoire issue d'une stratégie asynchrone satisfaisant (I)
- Supposons qu'il existe $t_* < t < t^*$ tq $x^{t_*} = x^{t^*} \neq x^t$
- Alors la trajectoire ne peut satisfaire (I)+(II)+(III) sur $[t_*, t^*]$

Pour qu'une trajectoire satisfaisant (I) revienne à un état déjà vu, il faut $\neg$(II) ou $\neg$(III)

Preuve technique...

((I) sur $\mathbb{N}$) + ((II)+(III) sur $[0, t]$) ⇒ Stabilisation...

- soit $(x^0, A^0) \to (x^1, A^1) \to (x^2, A^2) \to \ldots$ une trajectoire issue d'une stratégie asynchrone et équitable satisfaisant (I)
- Alors $\exists t$ tq
  la trajectoire satisfait (II) et (III) sur $[0, t]$, et

$$x^{t+k} = x^t \qquad \text{pour tout } k \in \mathbb{N}$$

Idée de Preuve :

- soit (t_i) une suite de temps tq tout neurone a été mis à jours entre t_i et t_{i+1}
- x_{t_i} boucle (espace fini) : $x_{t_k} = x_{t_l}$
- d'après le th. 3, $x_t = x_{t_k}$ pour tout $k \in [t_k, t_l]$
- la modification de la matrice ne change pas la comparaison au seuil

Quelques questions

1. Déterminer des classes de réseaux qui convergent avec Hebb mais pas sans Hebb.
2. Etude de la vitesse de convergence avec/sans la règle de Hebb
3. Peut-on jouer sur les paramètres de mise à jours des poids pour accélerer la convergence ?

Les deux idées générales :

- La règle de Hebb aide à la stabilisation (1)
- La règle de Hebb accélère la stabilisation (2,3)

L'introduction de la règle de Hebb ne peut pas produire de cycles limites, et au contraire elle les casse.

Quelques questions d'atteignabilité

Notation :

$\mathcal{A}_M(X_0, M_0) = \{M \text{ atteignable à partir de } (X_0, M_0)\}$

$\mathcal{A}_X(X_0, M_0) = \{X \text{ atteignable à partir de } (X_0, M_0)\}$

$\mathcal{F}_X(X_0, M_0) = \{X \text{ point fixe, atteignable à partir de } (X_0, M_0)\}$

1. Pour M_0 fixé, existe-t-il X_0 tq $\mathcal{A}_X(X_0, M_0) = \mathbb{B}^n$?
2. Pour M_0 fixé, existe-t-il X_0 tq $\mathcal{F}_X(X_0, M_0) = \mathbb{B}^n$?
3. Pour X_0 fixé, existe-t-il M_0 tq $\mathcal{A}_X(X_0, M_0) = \mathbb{B}^n$?
4. Pour X_0 fixé, existe-t-il M_0 tq $\mathcal{F}_X(X_0, M_0) = \mathbb{B}^n$?

Liens

- E. Goles, J. Olivos, Periodic behaviour of generalized threshold functions, Discr. Math. 30 :187-189, 1980.
  (convergence synchrone vers des points fixes ou cycles de longueurs 2)
- F. Robert, Les systèmes dynamiques discrets, vol. 19 of Mathematiques et Applications. Springer, 1995
  (matrice symétrique + diagonal non négative)
- M.-H. Shih, F.-S. Tsai, Growth dynamics of cell assemblies, SIAM J. Appl. Math., 69(4) :1110-1161, 2009.